

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : A01H 1/00, C07H 21/04, C07K 14/00, C12N 5/04, 5/10, C12P 19/34, C12Q 1/68	A1	(11) International Publication Number: WO 98/30083 (43) International Publication Date: 16 July 1998 (16.07.98)
(21) International Application Number: PCT/US98/00615 (22) International Filing Date: 9 January 1998 (09.01.98) (30) Priority Data: 08/781,734 10 January 1997 (10.01.97) US (71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612 (US). (72) Inventors: SHEN, Kathy; 44228 Country Club Drive, El Macero, CA 95618 (US). MEYERS, Blake; 904 Drake Drive, Davis, CA 95616 (US). MICHELMORE, Richard, W.; 36757 Russel Boulevard, Davis, CA 95616 (US). (74) Agents: EINHORN, Gregory, P. et al.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111 (US).		(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS (57) Abstract The present invention provides RG nucleic acids and proteins which confer disease resistance to plants. The nucleic acids can be used to produce transgenic plants resistant to pests. Antibodies to proteins of the invention are also provided.		

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS

5

The present application is a continuation-in-part application ("CIP") of U.S. Patent Application Serial No. ("USSN") 08/781,734, filed January 10, 1997. The
10 aforementioned application is explicitly incorporated herein by reference in its entirety and
for all purposes.

This invention was made with Government support under Grant Nos. 92-37300-7547 and 95-37300-1571, awarded by the United States Department of Agriculture.
15 The Government has certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates generally to plant molecular biology. In particular, it relates to nucleic acids and methods for conferring pest resistance in plants.
20 particularly lettuce.

BACKGROUND OF THE INVENTION

Recently, several resistance genes have been cloned by several groups from several plants. Many of these genes are sequence related. The derived amino acid
25 sequences of the most common class, *RPS2*, *RPM1* (bacterial resistances in *Arabidopsis* (Mindrinos *et al.* *Cell* 78:1089-1099 (1994)); Bent *et al.* *Science* 265:1856-1860 (1994); Grant *et al.*, *Science* 269:843-846 (1995)), *L6* (fungal resistance in flax; Lawrence, *et al.*, *The Plant Cell* 7:1195-1206 (1995)), and *N*, (virus resistance in tobacco; Whitham, *et al.*, *Cell* 78:1101-1115 (1994); and U.S. Patent No. 5,571,706), all contain leucine-rich
30 repeats (LRR) and nucleotide binding sites (NBS).

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

LRR domains can mediate protein-protein interactions and are found in a variety of proteins involved in signal transduction, cell adhesion and various other functions. LRRs are leucine rich regions often comprising 20-30 amino acid repeats where leucine and other aliphatic residues occur periodically. LRRs can function extracellularly or intracellularly.

Since the onset of civilization, plant diseases have had catastrophic effects on crops and the well-being of the human population. Plant diseases continue to effect enormous human and economic costs. An increasing human population and decreasing amounts of arable land make all approaches to preventing and treating plant pathogen destruction critical. The ability to control and enhance a plant's protective responses against pathogens would be of enormous benefit. Tissue-specific and temporal control of mechanisms responsible for plant cell death would also be of great practical and economic value. The present invention fulfills these and other needs.

What is needed in the art are plant disease resistance genes and means to create transgenic disease resistance plants, particularly in lettuce. Further, what is needed in the art is a means to DNA fingerprint cultivars and germplasm with respect to their disease resistance haplotypes for use in plant breeding programs. The present invention provides these and other advantages.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid constructs. These constructs comprise an RG (resistance gene) polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and an RG4 polypeptide. RG1, RG2, RG3, RG4, and the like, represent individual "RG families." Each "RG family," as defined herein, is a group of polypeptide sequences that have at least 60% amino acid sequence identity. Individual members of an RG family, *i.e.*, individual species of the genus, typically map to the same genomic locus. The invention provides for constructs comprising nucleotides encoding the RG families of the

invention, which can include sequences encoding a leucine rich region (LRR), and/or a nucleotide binding site (NBS), or both.

The invention provides for an isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence
5 identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide. In alternative embodiments, the nucleic acid construct comprises an RG polynucleotide which encodes an RG polypeptide comprising an leucine rich region (LRR), or, an RG polypeptide comprising a nucleotide binding site
10 (NBS). The nucleic acid construct can comprise a polynucleotide which is a full length gene. In another embodiment, the nucleic acid construct encodes a fusion protein.

In one embodiment, the nucleic acid construct comprises a sequence encoding an RG1 polypeptide. The RG1 polypeptide can be encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 and
15 SEQ ID NO:137 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).

In another embodiment, the nucleic acid construct comprises a sequence encoding an RG2 polypeptide. The RG2 polypeptide can be encoded by a polynucleotide
20 sequence selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87
25 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N);
30 SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In other embodiments, the nucleic acid construct comprises a RG3 sequence (SEQ ID NO:68) encoding an RG3 polypeptide (SEQ ID NO:138) (RG3). In other embodiments, the nucleic acid construct comprises an RG4 sequence (SEQ ID NO:69) encoding an RG4 polypeptide (SEQ ID NO:139) (RG4).

5 In other embodiments, the nucleic acid construct comprises a RG5 sequence (SEQ ID NO:134) encoding an RG5 polypeptide (SEQ ID NO:135). The RG5 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:134.

The invention also provides for a nucleic acid construct which comprises an RG7 sequence encoding an RG7 polypeptide. The RG7 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:136.

10 In further embodiments, the nucleic acid construct can further comprise a promoter operably linked to the RG polynucleotide. In alternative embodiments, the promoter can be a plant promoter; a disease resistance promoter; a lettuce promoter; a constitutive promoter; an inducible promoter; or, a tissue-specific promoter. The nucleic acid construct can comprise a promoter sequence from an RG gene linked to a heterologous polynucleotide.

The invention also provides for a transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide. The expression cassette can comprise a plant promoter or a viral promoter; the plant promoter can be a heterologous promoter. In one embodiment, the transgenic plant is lettuce. In alternative embodiments, the transgenic plant comprises an expression cassette which includes an RG polynucleotide selected from the group consisting of SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO:3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G); SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J); SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104

(RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W); SEQ ID NO:68 (RG3); SEQ ID NO:69 (RG4); SEQ ID NO:134 (RG5); or SEQ ID NO:136 (RG7).

The invention provide for a transgenic plant comprising an expression cassette comprising an RG polynucleotide which can encode an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W); an RG4 polypeptide as set forth by SEQ ID NO:72; an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135; or, an RG7 polypeptide.

The invention also provides for a method of enhancing disease resistance in a plant, the method comprising introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence. In this method, the plant can be a lettuce plant; and, the RG polynucleotide can encode an RG polypeptide selected from the group consisting of an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID

NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:72; SEQ ID NO:74; SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W). In this method, the promoter can be a plant disease resistance promoter, a tissue-specific promoter, a constitutive promoter, or an inducible promoter.

The invention also provides for a method of detecting RG resistance genes in a nucleic acid sample, the method comprising: contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and, wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample. In this method, the RG polynucleotide can be an RG1 polynucleotide, an RG2 polynucleotide, an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide. In this method, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide, and, the RG resistance gene can be amplified by the polymerase chain reaction. In one embodiment, the RG polynucleotide is labeled.

The invention further provides for an RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification, the figures and claims.

5 All publications, patents and patent applications cited herein are hereby expressly incorporated by reference for all purposes.

DETAILED DESCRIPTION OF THE INVENTION

This invention relates to families of RG genes, particularly from *Lactuca sativa*. Nucleic acid sequences of the present invention can be used to confer resistance in
10 plants to a variety of pests including viruses, fungi, nematodes, insects, and bacteria. Sequences from within the RG genes can be used to fingerprint cultivars or germplasm for the presence of desired resistance genes. Promoters of RG genes can be used to drive heterologous gene expression under conditions in which RG genes are expressed. Further, the present invention provides RG proteins and antibodies specifically reactive to RG
15 proteins. Antibodies to RG proteins can be used to detect the type and amount of RG protein expressed in a plant sample.

The present invention has use over a broad range of types of plants, including species from the genera *Cucurbita*, *Rosa*, *Vitis*, *Juglans*, *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*,
20 *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersicon*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Ciahorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Heterocallis*, *Nemesis*, *Pelargonium*, *Panieum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Pisum*, *Phaseolus*, *Lolium*, *Oryza*, *Zea*, *Avena*, *Hordeum*, *Secale*,
25 *Triticum*, and, *Sorghum*. In particularly preferred embodiments, species from the family *Compositae* and in particular the genus *Lactuca* are employed such as *L. sativa* and such subspecies as *crispa*, *longifolia*, and *asparagina*.

The nucleic acids of the present invention can be used in marker-aided selection. Marker-aided selection does not require the complete sequence of the gene or
30 precise knowledge of which sequence confers which specificity. Instead, partial sequences can be used as hybridization probes or as the basis for oligonucleotide primers to amplify nucleic acid, e.g., by PCR. Partial sequences can be used in other methods, such as to

follow the segregation of chromosome segments containing resistance genes in plants. Because the RG marker is the gene itself, there can be negligible recombination between the marker and the resistance phenotype. Thus, RG polynucleotides of the present invention provide an optimal means to DNA fingerprint cultivars and wild germplasm with respect to their disease resistance haplotypes. This can be used to indicate which germplasm accessions and cultivars carry the same resistance genes. At present, selection of plants (e.g., lettuce) for resistance to some diseases is slow and difficult. But linked markers allow indirect selection for such resistance genes. Moreover, RG markers also allow resistance genes to be identified and combined in a manner that would not otherwise be possible. Numerous accessions have been identified that provide resistance to all isolates of downy mildew (*Bremia lactucae*). However, without molecular markers it is impossible to combine such resistances from different sources. The nucleic acid sequences of the invention provide for a fast and convenient means to identify and combine resistances from different sources. The RG markers of the invention can also be used to identify recombinants that have new combinations of resistance genes in *cis* on the same chromosome.

In addition, RG markers may allow the identification of the Mendelian factors determining traits, such as field resistance to downy mildew. Once such markers have been identified, they will greatly increase the ease with which field resistance can be transferred between lines and combined with other resistances.

In another application, primers to RG sequences can be also designed to amplify sequences that are conserved in multiple RG family members. This gives genetic information on multiple RG family members. Alternatively, one or more primers can be made to sequences unique to a single resistance gene genus or a single RG specie. This allows an analysis of individual family groups (an RG genus) or an individual family member (a specie). Primers made to individual RGs at the edge of each cluster can be used to select for recombinants within the cluster. This minimizes the amount of linkage drag during introgression. Classical and molecular genetics has shown that pest resistance genes tend to be clustered in the genome. Pest resistance loci comprise arrays of genes and exhibit a variety of complex haplotypes rather than being simple alternate allelic forms. Pest resistance is conferred by families, or genuses, of related RG sequences, individual members, or species, of which have evolved to have a different specificity.

Oligonucleotide primers can be designed that amplify members from multiple haplotypes, or genuses, or amplify only members of one genus, or only amplify an individual specie. This will provide codominant information and allow heterozygotes to be distinguished from homozygotes.

5 Further, comparison of RG sequences will allow a determination of which sequences are critical for resistance and will ultimately lead to engineering resistance genes with new specificities. Resistance gene sequences were not previously available for lettuce. Marker-aided selection will greatly increase the precision and speed of breeding for disease resistance. Transgenic approaches will allow pyramiding of resistance genes
10 into a single Mendelian unit, transfer between sexually-incompatible species, substitute for conventional backcrossing procedures, and allow expression of other genes in parallel with resistance genes.

 The RG polynucleotides also have utility in the construction of disease resistant transgenic plants. This avoids lengthy and sometimes difficult backcrossing
15 programs currently necessary for introgression of resistance. It is also possible to transfer resistance polynucleotides between sexually-incompatible species, thereby greatly increasing the germplasm pool that can be used as a source of resistance genes. Cloning of multiple RG sequences in a single cassette will allow pyramiding of genes for resistance against multiple isolates of a single pathogen such as downy mildew or against multiple
20 pathogens. Once introduced, such a cassette can be manipulated by classical breeding methods as a single Mendelian unit.

 Transgenic plants of the present invention can also be constructed using an RG promoter. The promoter sequences from RG sequences of the invention can be used with RG genes or heterologous genes. Thus, RG promoters can be used to express a
25 variety of genes in the same temporal and spatial patterns and at similar levels to resistance genes.

Nucleic acids of the Invention and Their Preparation

RG Polynucleotide Families

30 The present invention provides isolated nucleic acid constructs which comprise an RG polynucleotide. In alternative embodiments, the RG polynucleotide is at least 18 nucleotides in length, typically at least 20, 25, or 30 nucleotides in length, more

typically at least 100 nucleotides in length, generally at least 200 nucleotides in length, preferably at least 300 nucleotides in length, more preferably at least 400 nucleotides in length, and most preferably at least 500 nucleotides in length.

In particularly preferred embodiments, the RG polynucleotide encodes a RG protein which confers resistance to plant pests. This RG protein can be longer, equivalent, or shorter than the RG protein encoded by an RG gene. In various embodiments, an RG polynucleotide can hybridize under stringent conditions to members of an RG family (an RG genus); *e.g.*, it can hybridize to a member of the RG1 RG family, such as an RG1 polynucleotide selected from the group consisting of: SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO: 3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G); SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J).

In other embodiments, the polynucleotide can also hybridize under stringent conditions to a member of the RG2 family; such as an RG2 polynucleotide selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In alternative embodiments, each RG2 gene can also include an AC15 sequence which hybridizes under stringent conditions to a polynucleotide selected from the group consisting of: SEQ ID NO:56 (AC15-2A); SEQ ID NO:57 (AC15-2B); SEQ ID NO:58 (AC15-2C); SEQ ID NO:59 (AC15-2D); SEQ ID NO:60 (AC15-2E); SEQ ID NO:61 (AC15-2G); SEQ ID NO:62 (AC15-2H); SEQ ID NO:63 (AC15-2I); SEQ ID

NO:64 (AC15-2J); SEQ ID NO:65 (AC15-2L); SEQ ID NO:66 (AC15-2N); SEQ ID NO:67 (AC15-2O).

In other embodiments, an RG polynucleotide can hybridize under stringent conditions to an RG3 (SEQ ID NO:68), an RG4 (SEQ ID NO:69), and RG5 (SEQ ID NO:135), and an RG7 (SEQ ID NO:137), RG family member.

The present invention further provides nucleic acid constructs which comprise an RG polynucleotide which encodes RG polypeptides from various RG families; such as an RG polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and RG4 polypeptide, and RG5 polypeptide, and an RG7 polypeptide.

Exemplary RG1 polypeptides have the sequences shown in SEQ ID NO:2 (RG1A), SEQ ID NO:4 (RG1B), SEQ ID NO:6 (RG1C), SEQ ID NO:8 (RG1D), SEQ ID NO:10 (RG1E), SEQ ID NO:12 (RG1F), SEQ ID NO:14 (RG1G), SEQ ID NO:16 (RG1H), SEQ ID NO:20 (RG1J). Exemplary RG2 polypeptides have the sequences shown in SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

An exemplary RG3 polypeptide has the sequence shown in SEQ ID NO:138. An exemplary RG4 polypeptide has the sequence shown in SEQ ID NO:139. RG polynucleotides will have at least 60% identity, more typically at least 65% identity, generally at least 70% identity, and preferably at least 75% identity, more preferably at least 80% identity, and most preferably at least 85%, 90%, or 95% identity at the deduced amino acid level. The regions where substantial identity is assessed can be inclusive or exclusive of the nucleotide binding site or the leucine rich region.

Vectors and Transcriptional Control Elements

The invention, providing methods and reagents for making novel species and genres of RG nucleic acids described herein, further provides methods and reagents for expressing these nucleic acids using novel expression cassettes, vectors, transgenic plants and animals, using constitutive and inducible transcriptional and translational *cis-* (e.g., promoters and enhancers) and *trans*-acting control elements.

The expression of natural, recombinant or synthetic plant disease resistance polypeptide-encoding or other (*i.e.*, antisense, ribozyme) nucleic acids can be achieved by operably linking the coding region a promoter (that can be plant-specific or not, constitutive or inducible), incorporating the construct into an expression cassette (such as an expression vector), and introducing the resultant construct into an *in vitro* reaction system or a suitable host cell or organism. Synthetic procedures may also be used. Typical expression systems contain, in addition to coding or antisense sequence, transcription and translation terminators, polyadenylation sequences, transcription and translation initiation sequences, and promoters useful for transcribing DNA into RNA. The expression systems optionally at least one independent terminator sequence, sequences permitting replication of the cassette *in vivo*, e.g., plants, eukaryotes, or prokaryotes, or a combination thereof, (e.g., shuttle vectors) and selection markers for the selected expression system, e.g., plant, prokaryotic or eukaryotic systems. To ensure proper polypeptide expression under varying conditions, a polyadenylation region at the 3'-end of the coding region can be included (see Li (1997) *Plant Physiol.* 115:321-325, for a review of the polyadenylation of RNA in plants). The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA (e.g., using *Agrobacterium tumefaciens* T-DNA replacement vectors, see e.g., Thykjaer (1997) *Plant Mol Biol.* 35:523-530; using a plasmid containing a gene of interest flanked by *Agrobacterium* T-DNA border repeat sequences; Hansen (1997) "T-strand integration in maize protoplasts after codelivery of a T-DNA substrate and virulence genes," *Proc. Natl. Acad. Sci. USA* 94:11726-11730.

To identify the promoters, the 5' portions of the clones described here are analyzed for sequences characteristic of promoter sequences. For instance, promoter sequence elements include the TATA box consensus sequence (TATAAT), which is usually 20 to 30 base pairs upstream of the transcription start site. In plants, further

upstream from the TATA box, at positions -80 to -100, there is typically a promoter element with a series of adenines surrounding the trinucleotide G (or T) N G (see, *e.g.*, Messing, in *Genetic Engineering in Plants*, pp. 221-227, Kosage, Meredith and Hollaender, eds. 1983). If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the RG coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from viral genes, such as T-DNA.

The nucleic acids of the invention can be expressed in expression cassettes, vectors or viruses which are transiently expressed in cells using, for example, episomal expression systems (*e.g.*, cauliflower mosaic virus (CaMV) viral RNA is generated in the nucleus by transcription of an episomal minichromosome containing supercoiled DNA, Covey (1990) *Proc. Natl. Acad. Sci. USA* 87:1633-1637). Alternatively, coding sequences can be inserted into the host cell genome becoming an integral part of the host chromosomal DNA.

Selection markers can be incorporated into expression cassettes and vectors to confer a selectable phenotype on transformed cells and sequences coding for episomal maintenance and replication such that integration into the host genome is not required. For example, the marker may encode biocide resistance, such as antibiotic resistance, particularly resistance to chloramphenicol, kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or Basta, to permit selection of those cells transformed with the desired DNA sequences, see for example, Blondelet-Rouault (1997) *Gene* 190:315-317; Aubrecht (1997) *J. Pharmacol. Exp. Ther.* 281:992-997. Because selectable marker genes conferring resistance to substrates like neomycin or hygromycin can only be utilized in tissue culture, chemoresistance genes are also used as selectable markers *in vitro* and *in vivo*. See also, Mengiste (1997) "High-efficiency transformation of *Arabidopsis thaliana* with a selectable marker gene regulated by the T-DNA 1' promoter," *Plant J.* 12:945-948, showing that the 1' promoter is an attractive alternative to the cauliflower mosaic virus (CaMV) 35S promoter for the generation of T-DNA insertion lines, the 1' promoter may be especially beneficial for the secondary transformation of transgenic strains containing the 35S promoter to exclude homology-mediated gene silencing.

The endogenous promoters from the RG genes of the present invention can be used to direct expression of the genes. These promoters can also be used to direct expression of heterologous structural genes. The promoters can be used, for example, in recombinant expression cassettes to drive expression of genes conferring resistance to any number of pathogens or pests, including fungi, bacteria, and the like.

Constitutive Promoters

In construction of recombinant expression cassettes, vectors, transgenics, of the invention, a promoter fragment can be employed to direct expression of the desired gene in all tissues of a plant or animal. Promoters that drive expression continuously under physiological conditions are referred to as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include those from viruses which infect plants, such as the cauliflower mosaic virus (CaMV) 35S transcription initiation region; the 1'- or 2'-promoter derived from T-DNA of *Agrobacterium tumefaciens*; the promoter of the tobacco mosaic virus; and, other transcription initiation regions from various plant genes known to those of skill. See also Holtorf (1995) "Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*," *Plant Mol. Biol.* 29:637-646.

Inducible Promoters

Alternatively, a plant promoter may direct expression of the plant disease resistance nucleic acid of the invention under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include pathogenic attack, anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters. For example, the invention incorporates the drought-inducible promoter of maize (Busk (1997) *supra*); the cold, drought, and high salt inducible promoter from potato (Kirch (1997) *Plant Mol. Biol.* 33:897-909).

Embodiments of the invention also incorporate use of plant promoters which are inducible upon injury or infection to express the invention's plant disease resistance (RG) polypeptides. Various embodiments include use of, *e.g.*, the promoter for a tobacco (*Nicotiana tabacum*) sesquiterpene cyclase gene (EAS4 promoter), which is expressed in wounded leaves, roots, and stem tissues, and upon infection with microbial pathogens (Yin

(1997) *Plant Physiol.* 115(2):437-451); the ORF13 promoter from *Agrobacterium rhizogenes* 8196, which is wound inducible in a limited area adjacent to the wound site (Hansen (1997) *Mol. Gen. Genet.* 254:337-343); the Shpx6b gene promoter, which is a plant peroxidase gene promoter induced by microbial pathogens (demonstrated using a fungal pathogen, see Curtis (1997) *Mol. Plant Microbe Interact.* 10:326-338); the wound-inducible gene promoter wun1, derived from potato (Siebertz (1989) *Plant Cell* 1:961-968); the wound-inducible *Agrobacterium pmas* gene (mannopine synthesis gene) promoter (Guevara-Garcia (1993) *Plant J.* 4:495-505).

Alternatively, plant promoters which are inducible upon exposure to plant hormones, such as auxins, are used to express the nucleic acids of the invention. For example, the invention can use the auxin-response elements E1 promoter fragment (AuxREs) in the soybean (*Glycine max* L.) (Liu (1997) *Plant Physiol.* 115:397-407); the auxin-responsive Arabidopsis GST6 promoter (also responsive to salicylic acid and hydrogen peroxide) (Chen (1996) *Plant J.* 10: 955-966); the auxin-inducible parC promoter from tobacco (Sakai (1996) 37:906-913); a plant biotin response element (Streit (1997) *Mol. Plant Microbe Interact.* 10:933-937); and, the promoter responsive to the stress hormone abscisic acid (Sheen (1996) *Science* 274:1900-1902).

Plant promoters which are inducible upon exposure to chemicals reagents which can be applied to the plant, such as herbicides or antibiotics, are also used to express the nucleic acids of the invention. For example, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, can be used (De Veylder (1997) *Plant Cell Physiol.* 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequence can be under the control of, *e.g.*, a tetracycline-inducible promoter, *e.g.*, as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) *Plant J.* 11:465-473); or, a salicylic acid-responsive element (Stange (1997) *Plant J.* 11:1315-1324. Using chemically- (*e.g.*, hormone- or pesticide-) induced promoters, harvesting of fruits and plant parts would be greatly facilitated. A chemical which can be applied to the transgenic plant in the field and induce expression of a polypeptide of the invention throughout all or most of the plant would make an environmentally safe defoliant or herbicide. Thus, the invention also provides for transgenic plants containing an inducible gene encoding for the RG

polypeptides of the invention whose host range is limited to target plant species, such as weeds or crops before, during or after harvesting.

Abcission promoters are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. In some embodiments, when a plant disease resistant polypeptide-encoding nucleic acid is under the control of such a promoter, rapid cell death, induced by expression of the invention's polypeptide, can accelerate and/or accentuate abcission, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like. Induction of rapid cell death at this time would accelerate separation of the fruit from the plant, greatly augmenting harvesting procedures. See, *e.g.*, Kalaitzis (1997) *Plant Physiol.* 113:1303-1308, discussing tomato leaf and flower abcission; Payton (1996) *Plant Mol. Biol.* 31:1227-1231, discussing ethylene receptor expression regulation during fruit ripening, flower senescence and abcission; Koehler (1996) *Plant Mol. Biol.* 31:595-606, discussing the gene promoter for a bean abcission cellulase; Kalaitzis (1995) *Plant Mol. Biol.* 28: 647-656, discussing cloning of a tomato polygalacturonase expressed in abcission; del Campillo (1996) *Plant Physiol.* 111:813-820, discussing pedicel breakstrength and cellulase gene expression during tomato flower abcission.

Tissue-Specific Promoters

Tissue specific promoters are transcriptional control elements that are only active in particular cells or tissues. Plant promoters which are active only in specific tissues or at specific times during plant development are used to express the nucleic acids of the invention. Examples of promoters under developmental control include promoters that initiate transcription only in certain tissues, such as leaves, roots, fruit, seeds, ovules, pollen, pistils, or flowers. Such promoters are referred to as "tissue specific". The operation of a promoter may also vary depending on its location in the genome. Thus, an inducible promoter may become fully or partially constitutive in certain locations.

For example, a seed-specific promoter directs expression in seed tissues. Such promoters may be, for example, ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed coat-specific, or some combination thereof. A leaf-specific promoter has been identified in maize, Busk (1997) *Plant J.* 11:1285-1295. The ORF13 promoter from *Agrobacterium rhizogenes* exhibits high activity in roots (Hansen (1997) *supra*). A maize pollen-specific promoter has been identified in maize (Guerrero (1990)

Mol. Gen. Genet. 224:161-168). A tomato promoter active during fruit ripening, senescence and abscission of leaves and, to a lesser extent, of flowers can be used (Blume (1997) *Plant J.* 12:731-746). A pistil specific promoter has been identified in the potato (*Solanum tuberosum* L.) SK2 gene, encoding a pistil-specific basic endochitinase (Ficker (1997) *Plant Mol. Biol.* 35:425-431). The Blec4 gene from pea (*Pisum sativum* cv. Alaska) is active in epidermal tissue of vegetative and floral shoot apices of transgenic alfalfa, making it a useful tool to target the expression of foreign genes to the epidermal layer of actively growing shoots. The activity of the Blec4 promoter in the epidermis of the shoot apex makes it particularly suitable for genetically engineering defense against insects and diseases that attack the growing shoot apex (Mandaci (1997) *Plant Mol Biol.* 34:961-965).

The invention also provides for use of tissue-specific plant promoters include a promoter from the ovule-specific *BEL1* gene described in Reiser (1995) *Cell* 83:735-742, GenBank No. U39944. Suitable seed specific promoters are derived from the following genes: *MAC1* from maize, Sheridan (1996) *Genetics* 142:1009-1020; *Cat3* from maize, GenBank No. L05934, Abler (1993) *Plant Mol. Biol.* 22:10131-1038; the gene encoding oleosin 18kD from maize, GenBank No. J05212, Lee (1994) *Plant Mol. Biol.* 26:1981-1987; viviparous-1 from *Arabidopsis*, Genbank No. U93215; the gene encoding oleosin from *Arabidopsis*, Genbank No. Z17657; *Atmyc1* from *Arabidopsis*, Urao (1996) *Plant Mol. Biol.* 32:571-576; the 2s seed storage protein gene family from *Arabidopsis*, Conceicao (1994) *Plant* 5:493-505; the gene encoding oleosin 20kD from *Brassica napus*, GenBank No. M63985; *napA* from *Brassica napus*, GenBank No. J02798, Josefsson (1987) *JBL* 26:12196-1301; the napin gene family from *Brassica napus*, Sjodahl (1995) *Planta* 197:264-271; the gene encoding the 2S storage protein from *Brassica napus*, Dasgupta (1993) *Gene* 133:301-302; the genes encoding oleosin a, Genbank No. U09118, and, oleosin B, Genbank No. U09119, from soybean; and, the gene encoding low molecular weight sulphur rich protein from soybean, Choi (1995) *Mol Gen, Genet.* 246:266-268. The tissue specific E8 promoter from tomato is particularly useful for directing gene expression so that a desired gene product is located in fruits. Other suitable promoters include those from genes encoding embryonic storage proteins.

One of skill will recognize that a tissue-specific promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, as

used herein a tissue-specific promoter is one that drives expression preferentially in the target tissue, but may also lead to some expression in other tissues as well.

The invention also provides for use of tissue-specific promoters derived from viruses which can include, *e.g.*, the tobamovirus subgenomic promoter (Kumagai (1995) *Proc. Natl. Acad. Sci. USA* 92:1679-1683; the rice tungro bacilliform virus (RTBV), which replicates only in phloem cells in infected rice plants, with its promoter which drives strong phloem-specific reporter gene expression; the cassava vein mosaic virus (CVMV) promoter, with highest activity in vascular elements, in leaf mesophyll cells, and in root tips (Verdaguer (1996) *Plant Mol. Biol.* 31:1129-1139).

In some embodiments, the nucleic acid construct will comprise a promoter functional in a specific plant cell, such as in a species of *Lactuca*, operably linked to an RG polynucleotide. Promoters useful in these embodiments include RG promoters. In additional embodiments, the nucleic acid construct will comprise a RG promoter operably linked to a heterologous polynucleotide. The heterologous polynucleotide is chosen to provide a plant with a desired phenotype. For example, the heterologous polynucleotide can be a structural gene which encodes a polypeptide which imparts a desired resistance phenotype. Alternatively, the heterologous polynucleotide may be a regulatory gene which might play a role in transcriptional and/or translational control to suppress, enhance, or otherwise modify the transcription and/or expression of an endogenous gene within the plant. The heterologous polynucleotide of the nucleic acid construct of the present invention can be expressed in either sense or anti-sense orientation as desired. It will be appreciated that control of gene expression in either sense or anti-sense orientation can have a direct impact on the observable plant characteristics.

Modifying and Inhibiting RG Gene Expression

The invention also provides for RG nucleic acid sequences which are complementary to the RG polypeptide-encoding sequences of the invention; *i.e.*, antisense RG nucleic acids. Antisense technology can be conveniently used to modify gene expression in plants. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the anti-sense strand of RNA will be transcribed. The construct is then transformed into plants and the antisense strand of RNA is produced. In plant cells, it has been shown that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, *e.g.*,

Sheehy (1988) *Proc. Nat. Acad. Sci. USA* 85:8805-8809; Hiatt et al., U.S. Patent No. 4,801,340.

Antisense sequences are capable of inhibiting the transport, splicing or transcription of RG-encoding genes. The inhibition can be effected through the targeting of genomic DNA or messenger RNA. The transcription or function of targeted nucleic acid can be inhibited, *e.g.*, by hybridization and/or cleavage. One particularly useful set of inhibitors provided by the present invention includes oligonucleotides which are able to either bind RG gene or message, in either case preventing or inhibiting the production or function of RG. The association can be through sequence specific hybridization. Such inhibitory nucleic acid sequences can, for example, be used to completely inhibit a plant disease resistance response. Another useful class of inhibitors includes oligonucleotides which cause inactivation or cleavage of RG message. The oligonucleotide can have enzyme activity which causes such cleavage, such as ribozymes. The oligonucleotide can be chemically modified or conjugated to an enzyme or composition capable of cleaving the complementary nucleic acid. One may screen a pool of many different such oligonucleotides for those with the desired activity.

Antisense Oligonucleotides

The invention provides for with antisense oligonucleotides capable of binding RG message which can inhibit RG activity by targeting mRNA. Strategies for designing antisense oligonucleotides are well described in the scientific and patent literature, and the skilled artisan can design such RG oligonucleotides using the novel reagents of the invention. In some situations, naturally occurring nucleic acids used as antisense oligonucleotides may need to be relatively long (18 to 40 nucleotides) and present at high concentrations. A wide variety of synthetic, non-naturally occurring nucleotide and nucleic acid analogues are known which can address this potential problem. For example, peptide nucleic acids (PNAs) containing non-ionic backbones, such as N-(2-aminoethyl) glycine units can be used. Antisense oligonucleotides having phosphorothioate linkages can also be used, as described in WO 97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197; Antisense Therapeutics, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Antisense oligonucleotides having synthetic DNA backbone analogues provided by the invention can also include phosphoro-dithioate, methylphosphonate,

phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, and morpholino carbamate nucleic acids, as described herein.

Combinatorial chemistry methodology can be used to create vast numbers of oligonucleotides that can be rapidly screened for specific oligonucleotides that have appropriate binding affinities and specificities toward any target, such as the sense and antisense RG sequences of the invention (for general background information, see, *e.g.*, Gold (1995) *J. of Biol. Chem.* 270:13581-13584).

Inhibitory Ribozymes

The invention provides for with ribozymes capable of binding RG message which can inhibit RG activity by targeting mRNA. Strategies for designing ribozymes and selecting the RG-specific antisense sequence for targeting are well described in the scientific and patent literature, and the skilled artisan can design such RG ribozymes using the novel reagents of the invention. Ribozymes act by binding to a target RNA through the target RNA binding portion of a ribozyme which is held in close proximity to an enzymatic portion of the RNA that cleaves the target RNA. Thus, the ribozyme recognizes and binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cleave and inactivate the target RNA. Cleavage of a target RNA in such a manner will destroy its ability to direct synthesis of an encoded protein if the cleavage occurs in the coding sequence, or, preventing transport of the message from the nucleus to the cytoplasm. After a ribozyme has bound and cleaved its RNA target, it is typically released from that RNA and so can bind and cleave new targets repeatedly.

Catalytic RNA molecules or ribozymes can also be used to inhibit expression of any plant gene. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs. The design and use of target RNA-specific ribozymes is described, *e.g.*, in Haseioff (1988) *Nature* 334:585-591.

In some circumstances, the enzymatic nature of a ribozyme can be advantageous over other technologies, such as antisense technology (where a nucleic acid

molecule simply binds to a nucleic acid target to block its transcription, translation or association with another molecule) as the effective concentration of ribozyme necessary to effect a therapeutic treatment can be lower than that of an antisense oligonucleotide. This potential advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single
5 ribozyme molecule is able to cleave many molecules of target RNA. In addition, a ribozyme is typically a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding, but also on the mechanism by which the molecule inhibits the expression of the RNA to which it binds. That is, the inhibition is caused by cleavage of the RNA target and so specificity is defined as the ratio
10 of the rate of cleavage of the targeted RNA over the rate of cleavage of non-targeted RNA. This cleavage mechanism is dependent upon factors additional to those involved in base pairing. Thus, the specificity of action of a ribozyme can be greater than that of antisense oligonucleotide binding the same RNA site.

The enzymatic ribozyme RNA molecule can be formed in a hammerhead
15 motif, but may also be formed in the motif of a hairpin, hepatitis delta virus, group I intron or RNaseP-like RNA (in association with an RNA guide sequence). Examples of such hammerhead motifs are described by Rossi (1992) *Aids Research and Human Retroviruses* 8:183; hairpin motifs by Hampel (1989) *Biochemistry* 28:4929, and Hampel (1990) *Nuc. Acids Res.* 18:299; the hepatitis delta virus motif by Perrotta (1992) *Biochemistry* 31:16;
20 the RNaseP motif by Guerrier-Takada (1983) *Cell* 35:849; and the group I intron by Cech U.S. Pat. No. 4,987,071. The recitation of these specific motifs is not intended to be limiting; those skilled in the art will recognize that an enzymatic RNA molecule of this invention has a specific substrate binding site complementary to one or more of the target gene RNA regions, and has nucleotide sequence within or surrounding that substrate
25 binding site which imparts an RNA cleaving activity to the molecule.

Sense Supression

Another method of suppression is sense suppression. Introduction of nucleic acid configured in the sense orientation has been shown to be an effective means by which to block the transcription of target genes. For an example of the use of this method
30 to modulate expression of endogenous genes see, Napoli et al., *The Plant Cell* 2:279-289 (1990), and U.S. Patent No. 5,034,323.

Cloning of RG Polypeptides

4

Synthesis and/or cloning of RG polynucleotides and isolated nucleic acid constructs of the present invention are provided by methods well known to those of ordinary skill in the art. Generally, the nomenclature and the laboratory procedures in recombinant DNA technology described below are those well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like are performed according to the manufacturer's specifications. These techniques and various other techniques are generally performed according to Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1989).

5

The isolation of RG genes may be accomplished by a number of techniques. For instance, oligonucleotide probes based on the sequences disclosed here can be used to identify the desired gene in a cDNA or genomic DNA library. To construct genomic libraries, large segments of genomic DNA are generated by random fragmentation, e.g. using restriction endonucleases, and are ligated with vector DNA to form concatemers that can be packaged into the appropriate vector. To prepare a cDNA library, mRNA is isolated from the desired organ, such as roots and a cDNA library which contains the RG gene transcript is prepared from the mRNA. Alternatively, cDNA may be prepared from mRNA extracted from other tissues in which RG genes or homologs are expressed.

15

The cDNA or genomic library can then be screened using a probe based upon the sequence of a cloned RG gene such as the genes disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species.

20

Those of skill in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent. As the conditions for hybridization become more stringent, there must be a greater degree of complementarity between the probe and the target for duplex formation to occur. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide. For example, the stringency of hybridization is conveniently varied by changing the polarity of the reactant solution through manipulation of the concentration of formamide within the range of 0% to 50%.

25

30

Alternatively, the RG nucleic acids of the invention can be amplified from nucleic acid samples using a variety of amplification techniques, such as polymerase chain reaction (PCR) technology, to amplify the sequences of the RG and related genes directly from genomic DNA, from cDNA, from genomic libraries or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Oligonucleotides can be used to identify and detect additional RG families and RG family species using a variety of hybridization techniques and conditions. Suitable amplification methods include, but are not limited to: polymerase chain reaction, PCR (PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, *ed.* Innis, Academic Press, N.Y. (1990) and PCR STRATEGIES (1995), *ed.* Innis, Academic Press, Inc., N.Y. (Innis)), ligase chain reaction (LCR) (Wu (1989) *Genomics* 4:560; Landegren (1988) *Science* 241:1077; Barringer (1990) *Gene* 89:117); transcription amplification (Kwoh (1989) *Proc. Natl. Acad. Sci. USA* 86:1173); and, self-sustained sequence replication (Guatelli (1990) *Proc. Natl. Acad. Sci. USA*, 87:1874); Q Beta replicase amplification and other RNA polymerase mediated techniques (*e.g.*, NASBA, Cingene, Mississauga, Ontario); see Berger (1987) *Methods Enzymol.* 152:307-316, Sambrook, and Ausubel, as well as Mullis (1987) U.S. Patent Nos. 4,683,195 and 4,683,202; Arnheim (1990) *C&EN* 36-47; Lomell *J. Clin. Chem.*, 35:1826 (1989); Van Brunt, *Biotechnology*, 8:291-294 (1990); Wu (1989) *Gene* 4:560; Sooknanan (1995) *Biotechnology* 13:563-564. Methods for cloning *in vitro* amplified nucleic acids are described in Wallace, U.S. Pat. No. 5,426,039.

The degree of complementarity (sequence identity) required for detectable binding will vary in accordance with the stringency of the hybridization medium and/or wash medium. The degree of complementarity will optimally be 100 percent; however, it should be understood that minor sequence variations in the probes and primers may be compensated for by reducing the stringency of the hybridization and/or wash medium as described earlier.

In some preferred embodiments, members of this class of pest resistance genes can be identified by their ability to be amplified by PCR primers based on the sequences disclosed here. Appropriate primers and probes for identifying RG sequences

from plant tissues are generated from comparisons of the sequences provided herein. See, e.g., Table 1. For a general overview of PCR see *PCR Protocols: A Guide to Methods and Applications*. (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), *Academic Press*, San Diego (1990), incorporated herein by reference.

5 Briefly, the first step of each cycle of the PCR involves the separation of the nucleic acid duplex formed by the primer extension. Once the strands are separated, the next step in PCR involves hybridizing the separated strands with primers that flank the target sequence. The primers are then extended to form complementary copies of the target strands. For successful PCR amplification, the primers are designed so that the
10 position at which each primer hybridizes along a duplex sequence is such that an extension product synthesized from one primer, when separated from the template (complement), serves as a template for the extension of the other primer. The cycle of denaturation, hybridization, and extension is repeated as many times as necessary to obtain the desired amount of amplified nucleic acid.

15 In the preferred embodiment of the PCR process, strand separation is achieved by heating the reaction to a sufficiently high temperature for an sufficient time to cause the denaturation of the duplex but not to cause an irreversible denaturation of the polymerase (see U.S. Patent No. 4,965,188). Template-dependent extension of primers in PCR is catalyzed by a polymerizing agent in the presence of adequate amounts of four
20 deoxyribonucleotide triphosphates (typically dATP, dGTP, dCTP, and dTTP) in a reaction medium comprised of the appropriate salts, metal cations, and pH buffering system. Suitable polymerizing agents are enzymes known to catalyze template-dependent DNA synthesis.

 Polynucleotides may also be synthesized by well-known techniques as
25 described in the technical literature. See, e.g., Carruthers *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams *et al.*, *J. Am. Chem. Soc.* 105:661 (1983). Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer
30 sequence.

RG Proteins

The present invention further provides isolated RG proteins encoded by the RG polynucleotides disclosed herein. One of skill will recognize that the nucleic acid encoding a functional RG protein need not have a sequence identical to the exemplified genes disclosed here. For example, because of codon degeneracy a large number of nucleic acid sequences can encode the same polypeptide. In addition, the polypeptides encoded by the RG genes, like other proteins, have different domains which perform different functions. Thus, the RG gene sequences need not be full length, so long as the desired functional domain of the protein is expressed.

The resistance proteins are at least 25 amino acid residues in length. Typically, the RG proteins are at least 50 amino acid residues, generally at least 100, preferably at least 150, more preferably at least 200 amino acids in length. In particularly preferred embodiments, the RG proteins are of sufficient length to provide resistance to pests when expressed in the desired plants. Generally then, the RG proteins will be the length encoded by an RG gene of the present invention. However, those of ordinary skill will appreciate that minor deletions, substitutions, or additions to an RG protein will typically yield a protein with pest resistance characteristics similar or identical to that of the full length sequence. Thus, full-length RG proteins modified by 1, 2, 3, 4, or 5 deletions, substitutions, or additions, generally provide an effective degree of pest resistance relative to the full-length protein.

The RG proteins which provide pest resistance will typically comprise at least one of an LRR or an NBS. Preferably, both are present. LRR and/or NBS regions present in the RG proteins of the present invention can be provided by RG genes of the present invention. In some embodiments, the LRR and/or NBS regions are obtained from other pest resistance genes. See, e.g., Yu *et al.*, *Proc. Natl. Acad. Sci. USA*, 93: 11751-11756 (1996); Bent *et al.*, *Science*, 265: 1856-1860 (1994).

Modified protein chains can also be readily designed utilizing various recombinant DNA techniques well known to those skilled in the art. For example, the chains can vary from the naturally occurring sequence at the primary structure level by amino acid substitutions, additions, deletions, and the like. Modification can also include swapping domains from the proteins of the invention with related domains from other pest resistance genes.

Pests that can be targeted by RG genes and proteins of the present invention include such bacterial pests as *Erwinia carotovora* and *Pseudomonas marginalis*. Fungal pests which can be targeted by the present invention include *Bremia lactucae*, *Marssonina panattoniana*, *Rhizoctonia solani*, *Olpidium brassicae*, root aphid, *Sclerotinia sclerotiorum* and *S. minor*, and *Botrytis cinerea* which causes gray mold. RG genes also provide resistance to viral diseases such as lettuce and turnip mosaic viruses.

Fusion Proteins

RG polypeptides can also be expressed as recombinant proteins with one or more additional polypeptide domains linked thereto to facilitate protein detection, purification, or other applications. Such detection and purification facilitating domains include, but are not limited to, metal chelating peptides such as polyhistidine tracts and histidine-tryptophan modules that allow purification on immobilized metals, protein domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequences such as Factor Xa or enterokinase (Invitrogen, San Diego CA) between the purification domain and plant disease resistant polypeptide may be useful to facilitate purification. One such expression vector provides for expression of a fusion protein comprising the sequence encoding a plant disease resistant polypeptide of the invention and nucleic acid sequence encoding six histidine residues followed by thioredoxin and an enterokinase cleavage site (*e.g.*, see Williams (1995) *Biochemistry* 34:1787-1797). The histidine residues facilitate detection and purification while the enterokinase cleavage site provides a means for purifying the desired protein(s) from the remainder of the fusion protein. Technology pertaining to vectors encoding fusion proteins and application of fusion proteins are well described, see *e.g.*, Kroll (1993) *DNA Cell. Biol.*, 12:441-53.

Antibodies Reactive to RG Polypeptides and Immunological Assays

The present invention also provides antibodies which specifically react with RG proteins of the present invention under immunologically reactive conditions. An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. "Immunologically reactive conditions" includes reference to conditions which allow an antibody, generated to a particular epitope of an antigen, to bind to that

epitope to a detectably greater degree than the antibody binds to substantially all other epitopes, generally at least two times above background binding, preferably at least five times above background. Immunologically reactive conditions are dependent upon the format of the antibody binding reaction and typically are those utilized in immunoassay protocols.

"Antibody" includes reference to an immunoglobulin molecule obtained by *in vitro* or *in vivo* generation of the humoral response, and includes both polyclonal and monoclonal antibodies. The term also includes genetically engineered forms such as chimeric antibodies (e.g., humanized murine antibodies), heteroconjugate antibodies (e.g., bispecific antibodies), and recombinant single chain Fv fragments (scFv). The term "antibody" also includes antigen binding forms of antibodies (e.g., Fab', F(ab')₂, Fab, Fv, rIgG, and, inverted IgG). See, Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL). An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. See, e.g., Huse *et al.* (1989) *Science* 246:1275-1281; and Ward, *et al.* (1989) *Nature* 341:544-546; and Vaughan *et al.* (1996) *Nature Biotechnology*, 14:309-314.

Many methods of making antibodies are known to persons of skill. A number of immunogens are used to produce antibodies specifically reactive to an isolated RG protein of the present invention under immunologically reactive conditions. An isolated recombinant, synthetic, or native RG protein of the present invention is the preferred immunogens (antigen) for the production of monoclonal or polyclonal antibodies.

The RG protein is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies can be generated for subsequent use in immunoassays to measure the presence and quantity of the RG protein. Methods of producing monoclonal or polyclonal antibodies are known to those of skill in the art. See, e.g., Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY.

Frequently, the RG proteins and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide

variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionucleotides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

The antibodies of the present invention can be used to screen plants for the expression of RG proteins of the present invention. The antibodies of this invention are also used for affinity chromatography in isolating RG protein.

The present invention further provides RG polypeptides that specifically bind, under immunologically reactive conditions, to an antibody generated against a defined immunogen, such as an immunogen consisting of the RG polypeptides of the present invention. Immunogens will generally be at least 10 contiguous amino acids from an RG polypeptide of the present invention. Optionally, immunogens can be from regions exclusive of the NBS and/or LRR regions of the RG polypeptides. Nucleic acids which encode such cross-reactive RG polypeptides are also provided by the present invention. The RG polypeptides can be isolated from any number plants as discussed earlier. Preferred are species from the family *Compositae* and in particular the genus *Lactuca* such as *L. sativa* and such subspecies as *crispa*, *longifolia*, and *asparagina*.

"Specifically binds" includes reference to the preferential association of a ligand, in whole or part, with a particular target molecule (i.e., "binding partner" or "binding moiety") relative to compositions lacking that target molecule. It is, of course, recognized that a certain degree of non-specific interaction may occur between a ligand and a non-target molecule. Nevertheless, specific binding, may be distinguished as mediated through specific recognition of the target molecule. Typically specific binding results in a much stronger association between the ligand and the target molecule than between the ligand and non-target molecule. Specific binding by an antibody to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. The affinity constant of the antibody binding site for its cognate monovalent antigen is at least 10^7 , usually at least 10^8 , preferably at least 10^9 , more preferably at least 10^{10} , and most preferably at least 10^{11} liters/mole. A variety of immunoassay formats are appropriate for selecting antibodies specifically reactive with a particular protein. For

example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically reactive with a protein. See Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific reactivity. The antibody may be polyclonal but preferably is monoclonal. Generally, antibodies cross-reactive to such proteins as RPS2, RPM1 (bacterial resistances in *Arabidopsis*, L6 (fungal resistance in flax, PRF (resistance to *Pseudomonas syringae* in tomato), and *N*, (virus resistance in tobacco), are removed by immunoabsorption.

Immunoassays in the competitive binding format are typically used for cross-reactivity determinations. For example, an immunogenic RG polypeptide is immobilized to a solid support. Polypeptides added to the assay compete with the binding of the antisera to the immobilized antigen. The ability of the above polypeptides to compete with the binding of the antisera to the immobilized RG polypeptide is compared to the immunogenic RG polypeptide. The percent cross-reactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% cross-reactivity with such proteins as RPS2, RPM1, L6, PRF, and *N*, are selected and pooled. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorption with these non-RG resistance proteins.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay to compare a second "target" polypeptide to the immunogenic polypeptide. In order to make this comparison, the two polypeptides are each assayed at a wide range of concentrations and the amount of each polypeptide required to inhibit 50% of the binding of the antisera to the immobilized protein is determined using standard techniques. If the amount of the target polypeptide required is less than twice the amount of the immunogenic polypeptide that is required, then the target polypeptide is said to specifically bind to an antibody generated to the immunogenic protein. As a final determination of specificity, the pooled antisera is fully immunosorbed with the immunogenic polypeptide until no binding to the polypeptide used in the immunosorption is detectable. The fully immunosorbed antisera is then tested for reactivity with the test polypeptide. If no reactivity is observed, then the test polypeptide is specifically bound by the antisera elicited by the immunogenic protein.

Production of transgenic plants of the invention

Isolated nucleic acid constructs prepared as described herein can be introduced into plants according techniques known in the art. In some embodiments, the introduced nucleic acid is used to provide RG gene expression and therefore pest resistance in desired plants. In some embodiments, RG promoters are used to drive expression of desired heterologous genes in plants. Finally, in some embodiments, the constructs can be used to suppress expression of a target endogenous gene, including RG genes.

To use isolated RG sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, for example, Weising *et al. Ann. Rev. Genet.* 22:421-477 (1988).

A DNA sequence coding for the desired RG polypeptide, for example a cDNA or a genomic sequence encoding a full length protein, will be used to construct a recombinant expression cassette which can be introduced into the desired plant. An expression cassette will typically comprise the RG polynucleotide operably linked to transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the RG gene in the intended tissues of the transformed plant.

Such DNA constructs may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG poration, particle bombardment and microinjection of plant cell protoplasts or embryogenic callus, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria.

Transformation techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski *et al. Embo J.* 3:2717-2722 (1984).

Electroporation techniques are described in Fromm *et al. Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein *et al. Nature* 327:70-73 (1987).

5 *Agrobacterium tumefaciens*-mediated transformation techniques are well described in the scientific literature. See, for example Horsch *et al. Science* 233:496-498 (1984), and Fraley *et al. Proc. Natl. Acad. Sci. USA* 80:4803 (1983). Although *Agrobacterium* is useful primarily in dicots, certain monocots can be transformed by *Agrobacterium*. For instance, *Agrobacterium* transformation of rice is described by Hiei *et al. Plant J.* 6:271-282 (1994). A particularly preferred means of transforming lettuce is described in Michelmore *et al., Plant Cell Reports*, 6:439-442 (1987).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired RG-controlled phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the RG nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans *et al., Protoplasts Isolation and Culture, Handbook of Plant Cell Culture*, pp. 124-176, Macmillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985.

15
20 Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al. Ann. Rev. of Plant Phys.* 38:467-486 (1987).

The methods of the present invention are particularly useful for incorporating the RG polynucleotides into transformed plants in ways and under circumstances which are not found naturally. In particular, the RG polypeptides may be expressed at times or in quantities which are not characteristic of natural plants.

25

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

30

Detection of RG Resistance Genes

The present invention further provides methods for detecting RG resistance genes in a nucleic acid sample suspected of comprising an RG resistance gene. The means by which the RG resistance gene is detected is not a critical aspect of the invention. For example, RG resistance genes can be detected by the presence of amplicons using RG resistance gene specific primers. Additionally, RG resistance genes can be detected by assaying for specific hybridization of an RG polynucleotide to an RG resistance gene. In some embodiments, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

In a typical detection method, the nucleic acid sample is contacted with an RG polynucleotide to form a hybridization complex. The hybridization complex may be detected directly (e.g., in Southern or northern blots), or indirectly (e.g., by subsequent primer extension during PCR amplification). The RG polynucleotide hybridizes under stringent conditions to an RG polynucleotide of the invention. Formation of the hybridization complex is directly or indirectly used to indicate the presence of the RG resistance gene in the nucleic acid sample.

Detection of the hybridization complex can be achieved using any number of well known methods. For example, the nucleic acid sample, or a portion thereof, may be assayed by hybridization formats including but not limited to, solution phase, solid phase, mixed phase, or *in situ* hybridization assays. Briefly, in solution (or liquid) phase hybridizations, both the target nucleic acid and the probe or primer are free to interact in the reaction mixture. In solid phase hybridization assays, probes or primers are typically linked to a solid support where they are available for hybridization with target nucleic in solution. In mixed phase, nucleic acid intermediates in solution hybridize to target nucleic acids in solution as well as to a nucleic acid linked to a solid support. In *in situ* hybridization, the target nucleic acid is liberated from its cellular surroundings in such as to be available for hybridization within the cell while preserving the cellular morphology for subsequent interpretation and analysis. The following articles provide an overview of the various hybridization assay formats: Singer *et al.*, *Biotechniques* 4(3):230-250 (1986); Haase *et al.*, *Methods in Virology*, Vol. VII, pp. 189-226 (1984); Wilkinson, "The theory and practice of *in situ* hybridization" In: *In situ Hybridization*, Ed. D.G. Wilkinson. IRL Press, Oxford University Press, Oxford; and *Nucleic Acid Hybridization: A Practical Approach*, Ed. Hames, B.D. and Higgins, S.J., IRL Press (1987).

The effect of the modification of RG gene expression can be measured by detection of increases or decreases in mRNA levels using, for instance, Northern blots. In addition, the phenotypic effects of gene expression can be detected by measuring nematode, fungal, bacterial, viral, or other pest resistance in plants. Suitable assays for determining pest resistance are well known. Micheltore and Crute, *Trans. Br. mycol. Soc.* 79(3): 542-546 (1982).

The means by which hybridization complexes are detected is not a critical aspect of the present invention and can be accomplished by any number of methods currently known or later developed. RG polynucleotides can be labeled by any one of several methods typically used to detect the presence of hybridized nucleic acids. One common method of detection is the use of autoradiography using probes labeled with ^3H , ^{125}I , ^{35}S , ^{14}C , or ^{32}P , or the like. The choice of radioactive isotope depends on research preferences due to ease of synthesis, stability, and half lives of the selected isotopes. Other labels include ligands which bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. Alternatively, probes can be conjugated directly with labels such as fluorophores, chemiluminescent agents or enzymes. The choice of label depends on sensitivity required, ease of conjugation with the probe, stability requirements, and available instrumentation. Labeling the RG polynucleotide is readily achieved such as by the use of labeled PCR primers.

The choice of label dictates the manner in which the label is bound to the probe. Radioactive probes are typically made using commercially available nucleotides containing the desired radioactive isotope. The radioactive nucleotides can be incorporated into probes, for example, by using DNA synthesizers, by nick translation with DNA polymerase I, by tailing radioactive DNA bases to the 3' end of probes with terminal deoxynucleotidyl transferase, by treating single-stranded M13 plasmids having specific inserts with the Klenow fragment of DNA polymerase in the presence of radioactive deoxynucleotides, dNTP, by transcribing from RNA templates using reverse transcriptase in the presence of radioactive deoxynucleotides, dNTP, or by transcribing RNA from vectors containing specific RNA viral promoters (e.g., SP6 promoter) using the corresponding RNA polymerase (e.g., SP6 RNA polymerase) in the presence of radioactive ribonucleotides rNTP.

The probes can be labeled using radioactive nucleotides in which the isotope resides as a part of the nucleotide molecule, or in which the radioactive component is attached to the nucleotide via a terminal hydroxyl group that has been esterified to a radioactive component such as inorganic acids, *e.g.*, ³²P phosphate or ¹⁴C organic acids, or esterified to provide a linking group to the label. Base analogs having nucleophilic linking groups, such as primary amino groups, can also be linked to a label.

Non-radioactive probes are often labeled by indirect means. For example, a ligand molecule is covalently bound to the probe. The ligand then binds to an anti-ligand molecule which is either inherently detectable or covalently bound to a detectable signal system, such as an enzyme, a fluorophore, or a chemiluminescent compound. Enzymes of interest as labels will primarily be hydrolases, such as phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescers include luciferin, and 2,3-dihydrophthalazinediones, *e.g.*, luminol. Ligands and anti-ligands may be varied widely. Where a ligand has a natural anti-ligand, namely ligands such as biotin, thyroxine, and cortisol, it can be used in conjunction with its labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

Probes can also be labeled by direct conjugation with a label. For example, cloned DNA probes have been coupled directly to horseradish peroxidase or alkaline phosphatase, (Renz. M., and Kurz, K. (1984) A Colorimetric Method for DNA Hybridization. *Nucl. Acids Res.* 12: 3435-3444) and synthetic oligonucleotides have been coupled directly with alkaline phosphatase (Jablonski, E., *et al.* (1986) Preparation of Oligodeoxynucleotide-Alkaline Phosphatase Conjugates and Their Use as Hybridization Probes. *Nuc. Acids. Res.* 14: 6115-6128; and Li P., *et al.* (1987) Enzyme-linked Synthetic Oligonucleotide probes: Non-Radioactive Detection of Enterotoxigenic *Escherichia Coli* in Faeca Specimens. *Nucl. Acids Res.* 15:5275-5287).

Definitions

Units, prefixes, and symbols can be denoted in their SI accepted form. Numeric ranges are inclusive of the numbers defining the range. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation, respectively. The

headings provided herein are not limitations of the various aspects or embodiments of the invention which can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the specification as a whole.

5 As used herein, the term "plant" includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds and plant cells and progeny of same. The class of plants which can be used in the methods of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants.

10 As used herein, "pest" includes, but is not limited to, viruses, fungi, nematodes, insects, and bacteria.

 As used herein, "heterologous" is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its original form. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form.

15 As used herein, "RG gene," alternatively referred to as "RLG gene," is a gene encoding resistance to plant pests, such as viruses, fungi, nematodes, insects, and bacteria, and which hybridizes under stringent conditions and/or has at least 60% sequence identity at the deduced amino acid level to the exemplified sequences provided herein. RG genes encode "RG polypeptides," alternatively referred to as "RLG polypeptides," which can comprise LRR motifs and/or NBS motifs. The RG polypeptides encoded by RG genes have at least 55% or 60% sequence identity, typically at least 65% sequence identity, preferably at least 70% sequence identity, often at least 75% sequence identity, more preferably at least 80% sequence identity, and most preferably at least 90% sequence identity at the deduced amino acid level relative to the exemplary RG sequences provided herein. The term "RG family" or "RG family genus" or "genus" includes reference to a group of RG polypeptide sequence species that have at least 60% amino acid sequence identity, and, the nucleic acids encoding these polypeptides. The individual species of a genus, i.e., the members of a family, typically are genetically mapped to the same locus.

25 As used herein, "RG polynucleotide" includes reference to a contiguous sequence from an RG gene of at least 18, 20, 25, 30, 40, or 50 nucleotides in length, up to

at least about 100 or at least about 200 nucleotides in length. In some embodiments, the polynucleotide is preferably at least 100 nucleotides in length, more preferably at least 200 nucleotides in length, most preferably at least 500 nucleotides in length. Thus, RG polynucleotide may be a RG gene or a subsequence thereof.

5 As used herein, "isolated," when referring to a molecule or composition, such as, for example, an RG polypeptide or nucleic acid, means that the molecule or composition is separated from at least one other compound, such as a protein, other nucleic acids (*e.g.*, RNAs), or other contaminants with which it is associated *in vivo* or in its naturally occurring state. Thus, an RG polypeptide or nucleic acid is considered isolated
10 when it has been isolated from any other component with which it is naturally associated, *e.g.*, cell membrane, as in a cell extract. An isolated composition can, however, also be substantially pure. An isolated composition can be in a homogeneous state and can be in a dry or an aqueous solution. Purity and homogeneity can be determined, for example, using analytical chemistry techniques such as polyacrylamide gel electrophoresis (SDS-
15 PAGE) or high performance liquid chromatography (HPLC).

 The term "nucleic acid" or "nucleic acid molecule" or "nucleic acid sequence" refers to a deoxyribonucleotide or ribonucleotide oligonucleotide in either single- or double-stranded form. The term encompasses nucleic acids, *i.e.*, oligonucleotides, containing known analogues of natural nucleotides which have similar or
20 improved binding properties, for the purposes desired, as the reference nucleic acid. The term also includes nucleic acids which are metabolized in a manner similar to naturally occurring nucleotides or at rates that are improved thereover for the purposes desired. The term also encompasses nucleic-acid-like structures with synthetic backbones. DNA backbone analogues provided by the invention include phosphodiester, phosphorothioate, phosphorodithioate, methylphosphonate, phosphoramidate, alkyl phosphotriester,
25 sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, morpholino carbamate, and peptide nucleic acids (PNAs); see *Oligonucleotides and Analogues, a Practical Approach*, edited by F. Eckstein, IRL Press at Oxford University Press (1991); *Antisense Strategies*, Annals of the New York Academy of Sciences, Volume 600, Eds. Baserga and Denhardt (NYAS 1992); Milligan (1993) *J. Med. Chem.* 36:1923-1937; *Antisense Research and Applications* (1993, CRC Press). PNAs contain non-ionic backbones, such
30 as N-(2-aminoethyl) glycine units. Phosphorothioate linkages are described in WO

97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197. Other synthetic backbones encompassed by the term include methyl-phosphonate linkages or alternating methylphosphonate and phosphodiester linkages (Strauss-Soukup (1997) *Biochemistry* 36:8692-8698), and benzylphosphonate linkages (Samstag (1996) *Antisense Nucleic Acid Drug Dev* 6:153-156). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide primer, probe and amplification product. Unless otherwise indicated, a particular nucleic acid sequence includes the complementary sequence thereof.

The term "exogenous nucleic acid" refers to a nucleic acid that has been isolated, synthesized, cloned, ligated, excised in conjunction with another nucleic acid, in a manner that is not found in nature, and/or introduced into and/or expressed in a cell or cellular environment other than or at levels or forms different than the cell or cellular environment in which said nucleic acid or protein is found in nature. The term encompasses both nucleic acids originally obtained from a different organism or cell type than the cell type in which it is expressed, and also nucleic acids that are obtained from the same cell line as the cell line in which it is expressed. invention.

The term "recombinant," when used with reference to a cell, or to the nucleic acid, protein or vector refers to a material, or a material corresponding to the natural or native form of the material, that has been modified by the introduction of a new moiety or alteration of an existing moiety, or is identical thereto but produced or derived from synthetic materials. For example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level, typically, under-expressed or not expressed at all. The term "recombinant means" encompasses all means of expressing, *i.e.*, transcription or translation of, an isolated and/or cloned nucleic acid *in vitro* or *in vivo*. For example, the term "recombinant means" encompasses techniques where a recombinant nucleic acid, such as a cDNA encoding a protein, is inserted into an expression vector, the vector is introduced into a cell and the cell expresses the protein. "Recombinant means" also encompass the ligation of nucleic acids having coding or promoter sequences from different sources into one vector for expression of a fusion protein, constitutive expression of a protein, or inducible expression of a protein, such as the plant disease resistant, or RG. polypeptides of the invention.

The term "specifically hybridizes" refers to a nucleic acid that hybridizes, duplexes or binds to a particular target DNA or RNA sequence. The target sequences can be present in a preparation of total cellular DNA or RNA. Proper annealing conditions depend, for example, upon a nucleic acid's, such as a probe's length, base composition, and the number of mismatches and their position on the probe, and can be readily determined empirically providing the appropriate reagents are available. For discussions of nucleic acid probe design and annealing conditions, see, *e.g.*, Sambrook and Ausubel.

The terms "stringent hybridization," "stringent conditions," or "specific hybridization conditions" refers to conditions under which an oligonucleotide (when used, for example, as a probe or primer) will hybridize to its target subsequence, such as an RG nucleic acid in an expression vector of the invention but not to a non-RG sequence. Stringent conditions are sequence-dependent. Thus, in one set of stringent conditions an oligonucleotide probe will hybridize to only one specie of the genus of RG nucleic acids of the invention. In another set of stringent conditions (less stringent) an oligonucleotide probe will hybridize to all species of the invention's genus but not to non-RG nucleic acids. Longer sequences hybridize specifically at higher temperatures. Stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium (if the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, *i.e.*, about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Often, high stringency wash conditions preceded by low stringency wash conditions to remove background probe signal. An example of medium stringency wash conditions for a duplex of, *e.g.*, more than 100 nucleotides, is 1x SSC at 45°C for 15 minutes (see Sambrook for a description of SSC buffer). An example low stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 4-6x SSC at 40°C for 15 minutes. a signal to noise ratio of 2x (or higher) than that observed for an unrelated

probe in the particular hybridization assay indicates detection of a "specific hybridization." Nucleic acids which do not hybridize to each other under stringent conditions can still be substantially identical if the polypeptides which they encode are substantially identical. This can occur, *e.g.*, when a nucleic acid is created that encodes for conservative
5 substitutions. Stringent hybridization and stringent hybridization wash conditions are different under different environmental parameters, such as for Southern and Northern hybridizations. An extensive guide to the hybridization of nucleic acids is found in, *e.g.*, Sambrook, Tijssen (1993) *supra*.

As used herein "operably linked" includes reference to a functional linkage
10 between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

15 In the expression of transgenes one of skill will recognize that the inserted polynucleotide sequence need not be identical and may be "substantially identical" to a sequence of the gene from which it was derived. As explained herein, these variants are specifically covered by this term.

In the case where the inserted polynucleotide sequence is transcribed and
20 translated to produce a functional RG polypeptide, one of skill will recognize that because of codon degeneracy, a number of polynucleotide sequences will encode the same polypeptide. These variants are specifically covered by the term "RG polynucleotide sequence". In addition, the term specifically includes those full length sequences substantially identical (determined as described herein) with an RG gene sequence which
25 encode proteins that retain the function of the RG protein. Thus, in the case of RG genes disclosed here, the term includes variant polynucleotide sequences which have substantial identity with the sequences disclosed here and which encode proteins capable of conferring resistance to nematodes, bacteria, viruses, fungi, insects or other pests on a transgenic plant comprising the sequence.

30 Two polynucleotides or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence, as described below. The term

"complementary to" is used herein to mean that the complementary sequence is identical to all or a specified contiguous portion of a reference polynucleotide sequence.

The terms "sequence identity," "sequence similarity" and "homology" refer to when two sequences, such as the nucleic acid and amino acid sequences or the polypeptides of the invention, when optimally aligned, as with, for example, the programs PILEUP, BLAST, GAP, FASTA or BESTFIT (see discussion, *supra*). "Percentage amino acid/nucleic acid sequence identity" refers to a comparison of the sequences of two polypeptides/nucleic acids which, when optimally aligned, have approximately the designated percentage of the same amino acids/nucleic acids, respectively. For example, "60% sequence identity" and "60% homology" refer to a comparison of the sequences of two RG nucleic acids or polypeptides which, when optimally aligned, have 60% identity. For example, in one embodiment, nucleic acids encoding RG polypeptides of the invention comprise a sequence with at least 50% nucleic acid sequence identity to SEQ ID NO:1. In other embodiments, the RG polypeptides of the invention are encoded by nucleic acids comprising a sequence with at least 50% sequence identity to SEQ ID NO:1, or, are encoded by nucleic acids comprising SEQ ID NO:1, or, have at least 60% amino acid sequence identity to the polypeptide of SEQ ID NO:2.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 55% or 60% sequence identity, generally at least 65%, preferably at least 70%, often at least 75%, more preferably at least 80% and most preferably at least 90%, compared to a reference sequence using the programs described above (preferably BESTFIT) using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding

identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 55% or 60%, preferably at least 70%, more preferably at least 80%, and most preferably at least 95%. Polypeptides having "sequence similarity" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under appropriate conditions. Appropriate conditions can be high or low stringency and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C to about 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent wash conditions are those in which the salt concentration is about 0.02 molar at pH 7 and the temperature is at least about 50°C. However, nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This may occur, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

Nucleic acids of the invention can be identified from a cDNA or genomic library prepared according to standard procedures and the nucleic acids disclosed here used as a probe. Thus, for example, stringent hybridization conditions will typically include at least one low stringency wash using 0.3 molar salt (*e.g.*, 2X SSC) at 65°C. The washes

are preferably followed by one or more subsequent washes using 0.03 molar salt (e.g., 0.2X SSC) at 50°C, usually 60°C, or mosre usually 65°C. Nucleic acid probes used to identify the nucleic acids are preferably at least 100 nucleotides in length.

As used herein, "nucleotide binding site" or "nucleotide binding domain" ("NBS") includes reference to highly conserved nucleotide-, *i.e.*, ATP/GTP-, binding domains, typically included in the "kinase domain" of kinase polypeptides, such as a kinase-1a, kinase 2, or a kinase 3a motif, as described herein. For example, the tobacco N and Arabidopsis RPS2 genes, among several recently cloned disease-resistance genes, share highly conserved NBS sequence. Kinase NBS subdomains further consist of three subdomain motifs: the P-loop, kinase-2, and kinase-3a subdomains (Yu (1996) *Proc. Acad. Sci. USA* 93:11751-11756). As discussed in detail herein, examples include the *Arabidopsis* RPP5 gene (Parker (1997) *supra*), the *A. thaliana* RPS2 gene (Mindrinos (1997) *supra*), and the flax L6 rust resistance gene (Lawrence (1995) *supra*) which all encode proteins containing an NBS; and Mindrinos (1994) *Cell* 78:1089-1099; and Shen (1993) *FEBS* 335:380-385. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can identify members having NBS domains, including any of the genus of NBS-containing plant disease resistant polypeptides of the invention.

As used herein, "leucine rich region" ("LRR") includes reference to a region that has a leucine content of at least 20% leucine or isoleucine, or 30% of the aliphatic residues: leucine, isoleucine, methionine, valine, and phenylalanine, and arranged with approximate repeated periodicity. The length of the repeat may vary in length but is generally about 20 to 30 amino acids. An LRR-containing polypeptide typically will have the canonical 24 amino acid leucine-rich repeat (LRR) sequence, which is present in different proteins that mediates molecular recognition and/or interaction processes; as described in Bent (1994) *Science* 265:1856-1860; Parker (1997) *Plant Cell* 9:879-894; Hong (1997) *Plant Physiol.* 113:1203-1212; Schmitz (1997) *Nucleic Acids Res.* 25:756-763; Hipskind (1996) *Mol. Plant Microbe Interact.* 9:819-825; Tornero (1996) *Plant J.* 10:315-330; Dixon (1996) *Cell* 84:451-459; Jones (1994) *Science* 266:789-793; Lawrence (1995) *Plant Cell* 7:1195-1206; Song (1995) *Science* 270:1804-1806; as discussed in further detail *supra*. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can

identify polypeptides having LRR domains, including any member of the genus of LRR-containing RG polypeptides of the invention.

The term "promoter" refers to a region or sequence determinants located upstream or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating and/or regulating transcription in plant cells; see also discussion on plant promoters, *supra*.

The term "constitutive promoter" refers to a promoter that initiates and helps control transcription in all tissues. Promoters that drive expression continuously under physiological conditions are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation; see also detailed discussion, *supra*.

The term "inducible promoter" refers to a promoter which directs transcription under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters; see also detailed discussion, *supra*.

The term "abscission-induced promoter" or "abscission promoter" refers to a class of promoters which are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. When the plant disease resistant polypeptide-encoding nucleic acid is under the control of an abscission promoter, rapid cell death, induced by expression of the invention's polypeptide, accelerates and/or accentuates abscission of the plant part, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like; see also detailed discussion, *supra*.

The term "tissue-specific promoter" refers to a class of transcriptional control elements that are only active in particular cells or tissues. Examples of plant promoters under developmental control include promoters that initiate transcription only (or primarily only) in certain tissues, such as roots, leaves, fruit, ovules, seeds, pollen, pistols, or flowers; see also detailed discussion, *supra*.

As used herein "recombinant" includes reference to a cell, or nucleic acid, or vector, that has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid to a form not native to that cell, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

As used herein, a "recombinant expression cassette" or "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements which permit transcription of a particular nucleic acid in a target cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of the expression vector includes a nucleic acid to be transcribed, and a promoter.

As used herein, "transgenic plant" includes reference to a plant modified by introduction of a heterologous polynucleotide. Generally, the heterologous polynucleotide is an RG structural or regulatory gene or subsequences thereof.

As used herein, "hybridization complex" includes reference to a duplex nucleic acid sequence formed by selective hybridization of two single-stranded nucleic acids with each other.

As used herein, "amplified" includes reference to an increase in the molarity of a specified sequence. Amplification methods include the polymerase chain reaction (PCR), the ligase chain reaction (LCR), the transcription-based amplification system (TAS), the self-sustained sequence replication system (SSR). A wide variety of cloning methods, host cells, and *in vitro* amplification methodologies are well-known to persons of skill.

As used herein, "nucleic acid sample" includes reference to a specimen suspected of comprising RG resistance genes. Such specimens are generally derived, directly or indirectly, from lettuce tissue.

The term "antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments or synthetic or recombinant analogues thereof which specifically bind and recognize analytes and antigens, such as a genus or subgenus of polypeptides of the invention, as described *supra*.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

5

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention.

Example 1

10 Example 1 describes the use of PCR to amplify RG genes from lettuce.

Multiple primers with low degeneracy, particularly at the 3' end, were designed based on the sequences of two known resistance genes from tobacco and flax.

DNA Templates

15 Lettuce genomic DNA was extracted from cultivar Diana and a mutant line derived from cultivar Diana using a standard CTAB protocol. To generate cDNA templates, RNA was isolated from cultivar Diana and the mutant following standard procedures; first strand cDNA was synthesized using Superscript reverse transcriptase from 1 Φ g total RNA as specified by the manufacturer (Life Technologies). BAC (bacterial artificial chromosome) clones from the *Dm3* region were isolated from a BAC library of
20 over 53,000 clones using marker AC15 that was known to be closely linked to *Dm3*. Bacterial plasmids containing clones of *L6* and *RPS2* were used as positive controls.

PCR with degenerate oligonucleotide primers

25 Oligonucleotide primers were designed based on conserved motifs in the nucleotide binding sites (NBS) of *L6*, *RPS2*, and *N*. Eight primers were made corresponding to the GVGKTT motif in the sense direction; each had 64-fold degeneracy. Six primers were made to the GLPLAL motif in the anti-sense direction; with either 16 or 256-fold degeneracy (Table 1).

30 Oligonucleotides included 14-mer adaptors of (CUA)₄ at the 5' end of the sense primers and (CAU)₄ at the 5' end of the antisense primers to allow rapid cloning of the PCR products into pAMP1 (Life Technologies).

PCR amplification was performed in 50 Φ l reaction volume with 1 Φ M of each of a pair of sense and antisense primers. The templates were denatured by heating to 94EC for 2 min. This was followed by 35 cycles of 30 sec at 94EC, 1 min at 50EC, 2 min at 72EC, with a single final extension of 5 min at 72EC. 25 ng of genomic DNA or cDNA was used. BAC clones as templates required less. The final dNTP concentration was 0.2 mM; $MgCl_2$ was 1.5 mM.

Forty-eight combinations of sense and antisense primers were tested on a panel of nine templates consisting of two genomic DNA samples, two cDNA preparations, three BAC clones and plasmids containing *L6* and *RPS2* as positive controls. Amplification from *L6* and *RPS2* resulted in fragments of 516 and 513 respectively. Seven combinations of primers resulted in fragments of approximately this size with multiple templates (Table 2). Primers that gave RLG products were: PLOOPAA, PLOOPAG, PLOOPGA, PLOOPGG, PLOOPAC, GLPL3, GLPL4.

15

20 (Intentionally left blank)

25

Table 1

DEGENERATE PRIMER SEQUENCES for NBS PCR

Sense primers based on GVGKTT amino acid sequence from L6, N and rps2 PLOOP motif:

PLOOPAG 5' GGN GTN GGN AAA ACG AC 3'

PLOOPAA 5' GGN GTN GGN AAA ACA AC 3'

PLOOPAT 5' GGN GTN GGN AAA ACT AC 3'

PLOOPAC 5' GGN GTN GGN AAA ACC AC 3'

PLOOPGG 5' GGN GTN GGN AAG ACG AC 3'

PLOOPGA 5' GGN GTN GGN AAG ACA AC 3'

PLOOPGT 5' GGN GTN GGN AAG ACT AC 3'

PLOOPGC 5' GGN GTN GGN AAG ACC AC 3'

Antisense primers based on GLPLAL amino acid sequence:

GLPL1 5' AGN GCN AGN GGN AGG CC 3'

GLPL2 5' AGN GCN AGN GGN AGA CC 3'

GLPL3 5' AGN GCN AGN GGN AGT CC 3'

GLPL4 5' AGN GCN AGN GGN AGC CC 3'

GLPL5 5' AAN GCC AAN GGC AAA CC 3'

GLPL6 5' AAN GCC AAN GGC AAT CC 3'

TABLE 2. Characteristics of RLGs isolated from lettuce.

	Template	Primers	Number ^a	Size ^b (bp)	Copy number ^c	Dm linkage
5	RLG1	genomic DNA	PLOOPGA+GLPL6	6/6	522	DM4,
		cDNA	PLOOPGA+GLPL6	1/5		DM13
		genomic DNA	PLOOPAA+GLPL6	5/5		
		cDNA	PLOOPAA+GLPL6	1/1		
	RLG2	BACH8	PLOOPGG+GLPL3	3/3	510	DM1, Dm3
	RLG3	genomic DNA	PLOOPGA+GLPL4	3/6	461	Dm5 Dm8
10	RLG4	genomic DNA	PLOOPGA+GLPL4	1/6	524	

^a Number of RLG sequences out of total number of clones sequenced.

^b Size of fragment amplified from the nucleotide binding domain.

^c Estimated copy number from genomic Southern blot analysis and numbers of clones in the BAC library.

Example 2

Example 2 describes the genetic analysis used to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes.

Bulked segregant analysis was performed to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes. DNA from individuals were pooled for each susceptible and resistant bulk. Amplified products were then mapped by RFLP analysis from our intraspecific mapping population. Resistances from four clusters of resistance genes as well as over six hundred markers have now been mapped on this population. Linkage analysis was done using JIONMAP or MAPMAKER mapping programs. Due to a suppression of recombination in the *Dm3* region, sequences were mapped relative to *Dm3* using a panel of deletion mutants that provided greater genetic resolution than the mapping population (Anderson *et al.* 1996). All blots were washed twice at 63EC in 2x SSC/1% SDS for 20 min, followed by one wash at 63EC in 1x SSC/0.1% SDS for 10 or 30 min.

Most of the RLG sequences were analyzed by bulked segregant analysis (BSA) using pools of resistant and susceptible individuals for each of the four clusters of resistance genes. In genomic Southern analyses, all the RLGs revealed numerous fragments of varying intensity. The numbers of bands was highly dependent of the stringency of hybridization. BSA demonstrated that RLG1 was linked to the *Dm4*, 7 and *Dm13* clusters. Segregation analysis confirmed this linkage.

RLG2 was derived from BAC H8 that was known to be from the *Dm3* region. BSA with RLG2 demonstrated that the polymorphic bands that distinguished the parents of our mapping population mapped to the *Dm1*, *Dm3* cluster. Several bands absolutely cosegregated with *Dm1* or *Dm3*. To provide finer genetic resolution, RLG2 was also mapped using a panel of *Dm3* deletion mutants. A number of fragments were missing in largest deletion mutant demonstrating that several RLG2 family members are physically located very close to *Dm3*. No fragment was missing in all deletion mutants; however, this is not unexpected as there is extensive duplication within the region.

Example 3

Example 3 describes the screening of a bacterial artificial chromosome library.

Over 53,000 BAC clones containing lettuce genomic DNA were screened with two of the amplified products. High density filters each containing 1536 clones were hybridized to ³²P labelled probes. Filters were washed at 65°C with 40 mM Na₂PO₄/0.1% SDS for 5 min followed by 20 min in the same solution.

To isolate additional RLG sequences we screened our genomic BAC library. Clones were identified that hybridized to RLG1 and RLG2. Nearly all the clones that hybridized to RLG2 also hybridized to marker AC15 that had already been shown by deletion mutant analysis to be clustered around *Dm3*. This provided further evidence for clustering of RLG2 sequences.

Using primers conserved within each family, part of the NBS was amplified from each unique BAC clone and sequenced. This revealed that members within each family varied from 64% identical at the deduced amino acid level. The most divergent members only weakly cross-hybridized to each other. Currently, RLG sequences are

considered to be part of the same family of sequences if they are at least 55% identical at the deduced amino acid level and map to the same region of the chromosome.

Example 4:

5 Example 4 describes the cloning, identification, sequencing and characterization of RG polynucleotide sequences; including use of RG sequences from plasmid and PCR products.

 Doubled stranded plasmid DNA clones and PCR products were sequenced using an ABI377 automated sequencer and fluorescently labelled di-deoxy terminators.
10 Sequences were assembled using Sequencher (Genecodes), DNASTar (DNASTar) and Genetics Computer Group (GCG, Madison, WI) software. Database searches were performed using BLASTX and FASTA (GCG) algorithms.

 Sequences flanking the NBS region for RLG2 and for some of RLG1 were obtained by a series of IPCR and the products sequenced directly. IPCR worked less well
15 for RLG1. Therefore RLG1 was subcloned from a BAC clone into pBSK (Stratagene) and the double stranded plasmid sequenced by long range sequencing.

 Initially, a total of 30 clones were sequenced. Three of these seven primer combinations yielded sequences that comprised continuous open reading frames with sequence identity to the NBS of known resistance genes. Seven out of 10 clones amplified
20 from genomic DNA with the primer pair PLOOPGA/GLP6 were 522 bp long; they were identical to each other and named RLG1. All six clones amplified from genomic DNA or cDNA using the primers PLOOPAA/GLP6 were similar/the same as RLG1. All three clones sequenced from BAC clone H8 were 510 bp long, identical to each other but different from RLG1 and were therefore designated RLG2. The 11 clones sequenced from
25 four other primer combinations had no similarity to any NBS motifs and therefore were not studied further. Therefore, sequencing resulted in the identification of clones containing NBS motifs representing four RLG sequences.

 Comparison of the deduced amino acid sequences of RLG1 and RLG2 to those of known resistance genes revealed that RLG1 and RLG2 are as similar to each other
30 as they are to resistance genes from other species and that this is the same level of identity shown between the known resistance genes (Table 3). The percent identity (upper quadrant) and percent identity (lower quadrant) were determined using the MEGALIGN

routine of the DNASTAR package. Identity refers to the proportion of identical amino acids; identity refers to the proportion of identical and similar amino acids and takes into account substitutions of amino acids with similar chemical characteristics. RG1 and RG2 are as similar to each other and to cloned resistance genes as cloned resistance genes from a variety of species are to each other. L6, resistance to *Melampsora lini* in flax (Lawrence *et al.*, 1995). N, resistance to tobacco mosaic virus in tobacco (Whitham *et al.*, 1994). PRF, required for resistance to *Pseudomonas syringae* in tomato. RPS2, resistance to *Pseudomonas syringae* in *Arabidopsis thaliana* (Bent *et al.*, 1994; Mindrinos *et al.*, 1994). RPM1, resistance to *Pseudomonas syringae* pv. *maculicola* in *A. thaliana* (Grant *et al.*, 1995). The initial RG1 and RG2, sequences were amplified from lettuce using degenerate primers.

Table 3**IDENTITIES OF****RESISTANCE GENE HOMOLOGUES**

		RG1	RG2	RG3	RG4	N gene	RPS2
Lettuce	RG1	***	22.7	15.0	29.2	25.4	23.8
Lettuce	RG2		***	32.2	21.6	22.7	33.0
Lettuce	RG3			***	17.2	15.0	32.8
Lettuce	RG4				***	44.3	22.7
Tobacco	N gene					***	21.6
<i>Arabidopsis</i>	RPS2						***

The regions homologous to the primers are included in this analysis as the genomic sequences for RLG1 and RLG2 were determined by IPCR. Interestingly, the genomic sequences for RLG1 exactly matched that of the primers used.

To obtain further evidence that we had amplified resistance genes, we amplified the regions flanking the NBSs of RLG1a and RLG2a by IPCR of BAC clones. These products were then directly sequenced without cloning to minimize the introduction of PCR artifacts. Sequence analysis of the 5' regions failed to detect any homology to known resistance genes. However, the sequence of the 3' region contained leucine-rich

repeats (LRRs). When this sequence was used to search GENBANK using BLASTX, it detected identity to the *Arabidopsis* resistance gene, *RPS2*. This region does not contain as regular LRRs as in some resistance genes; however, the repeat structure seems to be consistent with that of the flax resistance gene, *L6*. Therefore, the presence of an LRR region is further evidence that the sequences we amplified using degenerate oligonucleotide primers are probably resistance genes.

The sequences of the IPCR products also provided the genomic sequences of the regions complementary to the sequences of the degenerate oligonucleotide primers.

The genomic sequences for RLG1 were identical to one of the primers in the mixture.

The RLG sequences are resistance genes as supported by three criteria: the presence of multiple sequence motifs characteristic of resistance genes, genetic cosegregation with known resistance genes, and their existence as clustered multi-gene families. The presence of LRR regions in a similar position relative to the NBS as in cloned resistance genes provides stronger evidence than relying solely sequence similarity between NBS regions.

The clustering of RLG sequences at the same position as the known clusters of resistance genes make them strong candidates for encoding resistance genes. The hybridization patterns and genetic distribution of the RLG sequences are similar to that of cloned resistance genes in other species. Most of these hybridize to small multigene families and preliminary genetic evidence indicates that they are clustered in the genome. Therefore, the degenerate primers that we designed from other resistance genes seemed to have been specific enough to amplify resistance genes rather than P-loop containing proteins in general.

(intentionally left blank)

SEQ ID NO: 1

RLG1A
[Strand]

```

1   ATCGTAACCGTTCGTACGAG  ANCGCTGKCCCTCTTCATC  TTTTGTCATATGTCATATTC  TCATNIAITMTGCCACATNT
81  AATTTTGTGGTTATTTTAA  TTAATTTTATTCACATGT  CATTTTATGAGTTTTCAT  TTTATTGAGTTTCACATAAT
161 ATTTAAATGTAATAACAATA  AATGCATATTTATTTTCTT  TAAATAAACGCATATAATAT  ATAGATTAAAAATCATATAAT
241 ACATAGGTTAAACATCATA  ATACATATGTTTCATCCCCAG  TTTATTTATATGTCATCACC  TTAATTTATTTATTTATTTAT
321 TTATTAGAGTAGATCTTT  TGATATTAATAAATTTAAT  TTGTTCAAATTTAAATTA  TTAATAATCCCAATTTGA
401 ATAAAAATTAAAAAATGGN  CCCACCATTAGTCCATCACT  TTTTCAGCTCATCAATATCG  TGAGTATTCTCCCTCGTTTC
481 CACCCCAATCAATATTTCCA  GCGAATGACAGACTCCCTACG  GCGTTTCTGAATTTGCGTTC  CGACACTGTTCATTGAAGGA
561 GATAATAAATCAAAATGGAG  TGCTCCAAATGTTCAATGCTG  ATGAAAGGTGAATTTGATGT  GAAGANAATGTCAGCGATCN
641 ATCTCCATCCGGAACCCACC  ACATTATCAGTGTACCACCA  AACCACTCAAACCGGYGGAA  GTAGRRAKACWRKAAAGTCA
721 TGAAGAATAGATTATTTTIG  TCCTCATGGGCTGACTGAGG  AGCGGGTTTAGTTCATCATT  TTTCTTTGANCAAAGAATTA
801 TCGGTCCATCGAATTTTAC  ATCGACAAAGAAGTTTCACT  TCGCAATGTTTGTAAACA  ATTTTAAATCTTTTATCTT
881 TTCGTTGAACTCCTCAATT  GCAACTTGCAACTTGCAACT  TTTGGGCCCACTAATTTGTG  GTGGCGTTAATTTAATCCA
961 CATATTCACGTAAACAATA  ATTCAAATCGATCTCTGTTC  ATCCAATTCATCAACATCTC  TTGATAATTGAATCATTTCA
1041 CGCTTCATCCATTTCATCCA  CATCTATACATATTTCTGTC  CTCTATCATATTAACGAT  GGCTGAAATCGTTCTTTCTG
1121 CCTTCCTTGACAGTGGTGT  GAAAAGCTGGCATYTGAAGC  CTTGAAGAAGATTGTTGCGT  CCAAAGAATTGAATCTGAG
1201 CTTAAGAAATGAAGGAGAC  ATTAGACCAAATCCAAGATC  TGCTTAACGATGCTTCCAG  AAGGAAGTAACTAATGAAGC
1281 CGTTAAAGATGGCTGAATG  ATCTCCAACATTTGGCTTAT  GACTAGACGACCTACTTGA  TGATTTGCAACTGAAGCTG
1361 TTCACWCTGAGTTGACCGAG  GAGGGTGGAGCCTCCTCCAG  TATGGTAAGAAAATAATCC  CAAGTTGTTGCACAAGTTTC
1441 TCACAAAGTAATAGGATGCA  TGCCAAGTTAGATGATATG  CCACCAGGTTACAGAACTG  GTAGAGGCAAAAATAATCT
1521 TGGTTTAAAGTGTGATAACAT  ATGAAAGCCAAAATTTGAA  AGGTATGAGGCGCTCTTGGT  AGATGAAAGCGGTACTGTCG
1601 GACGTGAAGATGATAAGAAA  AAATTGCTGGAGAAGCTGTT  GGGGATAAAGATGAATCAG  GGAGTCAAACCTTCAGCATC
1681 GTGCCCATAGTTGGTATGGG  TGGAGTTGGTAAACAACATC  TAGCTAGACTTTTGTATGAT  GAAAAGAAAGTGAAGGATCA
1761 CTTGCAACTCAGGGCTTGGG  TTTGTGTTTCTGATGAGTTC  AGTGTCCCAATATAAGCAG  AGTTATTTATCAATCTGTGA
1841 CTGGGGAAAAGAAGGAGTTT  GAAGACTTAAATCTGCTTCA  AGAAGCTCTTAAAGAGAAAC  TTAGGAACCAGCTATTTCCTA
1921 ATAGTTTTGGATGATGTGTG  GTCTGAAAGCTATGTTGAT  GGGAGAAATTAGTGGGCCCA  TTCTTGGCGGGTCTCCCTGG
2001 AAGTAGAATAATCATGACAA  CTCGGAAGGAGCAATTTGCT  AGAAAGCTGGGCTTTTCTCA  TCAAGACCCCTCTGGAGGGTC
2081 TATCACAAGATGATGCTTIG  TCCTTGTGTTGCTCAACACGC  ATTTGGTGTACCAAACTTTG  ATTCAATCCAACCTAAGG
2161 CCACATGGAGAACTGTTTGT  GAAGAAATGTGATGGCTTAC  CTCTAGCTTTAAGAACACTT  GGAAGGTTATTAAAGACAAA
2241 AACAGACAGGGAACATGGA  AGGAGCTGTTGGATAGTGAG  ATATGGAGGTTAGGAAAGAG  CGATGAGATTGTTCCGGCTC
2321 TTAGCAATAGCTACATGAT  CTTTCTGCCWCTTTGAAGCT  RTTTRTTGCAATAYTGTCTCT  TGTTTCCCAAGCTATGAG
2401 TTTGACAAGGAGGAGTTGAT  TCTATTGTGGATGGCAGAAG  GGTTTTTCACCAACCAACT  AYAAACAAGTCAAGCAACG
2481 KTTGGGCTCTGAAATATTTT  AAGAGTTTGTGCAAGRTCR  TTTTTCACATGCTCCTTAA  TRRCAATCSTTGTGTTGA
2561 TGCATGACCTAATGAATGAT  TTGGCTACATTGTTGCTGG  AGAATTTTTTCAAGGTTAG  ACATAGAGATGAAGAAGGAA
2641 TTTAGGATGSAATCTTTTGA  RAAGCACCGMCATATGTCAT  TTGTATGTGAGRATTACATA  GGTACAAAARGTTTCGAGCC
2721 ATTTAGAGGAGCTAAAAATT  TGAGAACATTTTACCATTC  TCTGTTGGGGTGGTAGAAGA  TTGGAAGATGTTTACTTTAT
2801 CAAACAAGGCTTTGAATGAC  WTACTTCARGATTTACCATT  GTTAAGGGTCTTRAKTTTGA  TTRRTCTTAYAAATAASYRAG
2881 GTACCAARAATCTGSGTAG  TATGAASCATTTGCGGTATC  TTAATCTATCWGRAACTTWA  ATCACMCATTTACCGGAWA
2961 TKTCCTGCAATCTTTATAATT  TACARACCTGATTTGTCTCT  GGCTGTGAMTATTTAGTTAA  KTTGCCCAARACCTTCTCAA
3041 ASCTTAAAAATTTGCASCAT  TTTGACATGAGGGRTACTCC  KAAKTTAARAACATGCCCT  TARGGATGGTGARTTGAAA
3121 ARTCTACAAACTCTCTTNG  TAACATTGGCATAGCAATAA  CCGAGCTTAAGAACTTGCAM  AAYCTCCATGGGAAARTTTG
3201 TATTGGCGGGCTGGGAAAAA  TGGAAATGCMGTGKGATGC  ACGTTAAGCGAACTTGTCTC  AAAAAAGGTTWAATGARTTA
3281 NAAACTGGRTWTTGGGGTGA  TFAATTTAATGTTTTCCGAA  ATGGGAACACTTGAAAAAGA  AGTCTCTCAATGAAGTATGC
3361 CTCATAATGGTACTCTANAA  AAAACCCANAAATATGTCTA  TAGGGGTTATAGAGTTTCCA  AATTGGGTTGGTTNCACTAA
3441 GGGTTTTCTGAACTAGAGAT  GTGTTCATGGTGTATGAAAA  AGANTGTTTTACGTAGTTTC  ATCAATCACCAGTGGGAAA
3521 TAGATGATATTTTCAGGGCY  TACTGATGAGATGTGGAGAG  GTATGATAGGGTNTCTTGGG  GCGGTAGAAGAAATAAGCAT
3601 CCATCTCTGTGAATGAAATAA  GATATYTGTTGGGAATCAGAA  GCAGAGGCAAGTAAGGTTC  TATGAATTTAAAGAAGTTGG
3681 ATTTAGGTGAATGTGAAAT  TTGGTGAGTTTAGGGGAGAA  AAAGGAGGATAATCATAATA  TTAATAGTGGGAGCAGCCTA
3761 ACATCTTTTAGGAGGTTGAA  TGTATGGAGATGTAACAGCT  TGGAGCATTGCAAGGTGTCCA  GATAGCATGGAGAATTTGTA
3841 TATGCAATGTGTGATTCAA  THACATCCGCTCTCTTCCCA  ACAGGAGGAGGACAGAAGAT  CAAGTCAATACCATCACTG
3921 ATTGCAAGAAGCTTTCCGAA  GAGGAGTTGGGAGGACGAGA  GAGGACAAGAGTCTTATAA  ACTCAAAAATGCAGATGCTT
4001 GAATCTAGATATACGTAA  TTGGCCAAATCTGAAATCTA  TCAGTGAATTAGTGTCTTC  ATTCACCTGAACAGATTATA
4081 TATATCAACTGTCCGAGTR  TGGAGTCAATTTCTGACCAT  GAGTTGCCAAATCTCACCTC  CTTACAGATCGAAGGAGAG
4161 GACAGCGATTTTCTGACGAA  CGGTTACGATTCGACTGGCC  GTGCTTTT

```

SEQ ID NO: 2

RLG1B

[Strand]

```
1  AACCGTTCGT ACCGAAATCG CTGTCTCTCT CTTCCTGTAA TATAATGATA AGAAAAATA TGATTAAAGG
71  TTTAAATCCA AATTCATTTA TTCCACCGGT GATATGATGC ACTAGCTGTA GTATGCAAAA ACAGTATTAT
141 AAATGCTAAC CAAAACAGCA GCTAAGAAAC AATATAAATA ATGGTTTGAA TCGTCTTTTC TCCGTACAQT
211 CATTTCTTCC AAATCCCTAT CATTCATACA TACAAGTGCT CCCATATTAG GTTTTCACTA TAAGCAATGG
281 CTGAAATCCT TGGTCTGCGG TTCCTTTCGG TGTCTTTTGA AAAGCTTGCT TCTGAAGCCT TGAAGAGGGT
351 TGCTTGCTCC AAAGTAATTG ACAAGGAGCT CGAGAAATTG AATAGCTCAT GAATCAATAT AAAAGCTCTG
421 CTCAATGATG CTTCTCAGAA GGAATAAAGT AAGGAAGCTG TTAAGAATG GTTGAATGCT CTTCAACATT
491 TGCTTACGA CATAGATGAT CTACTTGGCG ATTTGGCAAC CAAAGCTATC CATCGTAAGT TCTCTGAGGA
561 ATACGGGGCC ACCATCAACA AGGTACGAAA GTTAAATCCA TCTGTCTTCT CTAGTTTGTC AAGTACTAAG
631 ATGCGCAACA AGATACATAA TATTACCAGC AAGTTACAAG AACTATTAGA AGAGAGAAAT AATCTTGGAT
701 TATGTGAAT TGGTGAAGC CGAAAACCTC GAAATAGAAA ATCAGAGACC TCCTGTCTAG ATCCATCTAG
771 TATTGTTGGA CGCAGATG ATAAGGAAGC GTTGCTTCTC AAGCTATATG AACCATGTGA TAGAACTTT
841 AGCATCTTGC CNATAGTTGG TATGGGTGGG TTAGATAAGA CCACTTTAGG TAGACTTTTG TATGATNAAA
911 TGCAAGTGAA GGATCACTTC GAACTCAAGG CGTGGGTTTG TGTCTCTGAT GAGTTTGATA TCTTCGGTAT
981 AAGCAAAACC ATTTTCGAAT CGATAGAGGG GGGAAACCAA GAGTTTAAGG ATTTAAATCT GCTTCAGGTG
1051 GCTTTAAAGG AGAAAATCTC AAAGAAACGA TTCTTGTGTG TTCTTGATGA TGTATGGAGC GAGAGCTATA
1121 CTGATTGGGA AATCTAGAA CGTCCATTTC TAGCAGGAGC ACCAGGAAGT AAAGTAATCA TCACAACCCG
1191 CAAGTCTGCG TTGCTAAACC AATTGGGTCA TGATCAACCA TACCAATTGT CTGATTTGTC ACATGACAAT
1261 GCTCTATCCT TATTTTGTC ACACGCATTT GGTGTAATA GCCTTGATTC ACATCCGATA CTTAAACCCAC
1331 ATGGTGAAGG TATTGTTGAA AAATGTGATG GTTTGCCATT GGCTTTGATT GCACCTGGGA GGTATTTGAG
1401 GACAAAAGA GATGAGGAAG AATGGAAGGA ACTATTGAAT AGTGAGATAT GGAGGTTAGG AAAGAGAGAT
1471 GAGATCTATC CGGTCTTAG ACTAAGCTAT AATGATCTTT CTGCTCTTTT GAAGCAGTTG TTTGCAATTT
1541 GCTCCTTGTT CCCCAAGAC TATGTGTTCA ACAAGGAGAA GTTGATTTTA TTATGGATGG CAGAAGGGTT
1611 TTTGCACAAT GAAAATACAA ACAAGTCAAT GGAACGCTTA GNTCTTGAAT ATTTTGAAGA CTTGTTGTCA
1681 AGGTCTTTT TTCAACATGC ACTCGATGAC AAATCGTTGT TTGTGGTGCA CGACCTCATG AATGACTTGG
1751 CCACATCTGT TGCTGGAGAT TATTTTTTAA GATTAGACAT TGAAATGAAA AAGGAAGCTT TGGAAAAATA
1821 CCGACATATG TCATTTGTTT GTGAGAGTTA CATGGTTTAC AAAAGGTTTG AACCATTTAA AGGAGCTAAA
1891 AAATTGAGAA CTTTCTTAGC AATGCCGTGT GGGATGATAA AAAGTTGGAC AACATTTTAC TTATCAAATA
1961 AGGTCTTTGA TGACTTACTT CACGAATTAC CATTGTTGAG AGTTCTAAGT TTGAGTTATC TTAGCATCAA
2031 GGAGGTACCT GAAATAATAG GCAATTTGAA ACTCTGCGG TATCTTAATT TATCACACAC GAGTATCACA
2101 CATTTTACCAG AAAATGTCTG CAATCTTTAC AACTTACAAA CATTGATCCT TTGTGGCTGT TGTMTTATAA
2171 CCAAGTTTCC CACCAACTTC TTAAAGCTTA GAAATTTACG GCATTTGGAC ATTAGCGATA CTCCCGGTTT
2241 GAAGAAGATG TCCTCGGGGA TTGGTGAATT GAAGAACCTA CACACYCTCT CCAAGCTCAT TATTTGAGGT
2311 GAAAAATAGAC TAAACGAGCT TAAGAACTTA CAAAATCTCC ATG
```

RLG1b - Diana
[Strand]

```
1  TACTACTACT AGAATTCGGT GTTGTAAGA CGACTCTAGC TAGACTTTTG TATGAGGAAA TGCAAGGGAA
71  GGATCACTTC GAACTTAAGG CGTGGGTATG TGTTCCTGAT GAGTTTGATA TCTTCAATAT AAGCAAAATT
141 ATCTTACAAT CGATAGGTGG TGGAAACCAA GAATTTACGG ACTTAAACCT GCTTCGAGTA GCTTTAAABAG
211 AGAAGATCTC AAAGAAAAGa TTTCTCTCTG TTCTTGATGA TGTTCGGAGT GAAAGCTATA CCGATTGGGA
281 AATTNTAGAA CGCCCATTTT TTGCAGGGGC ACCTGGAAGT AAGATTATTA TCACCACCCG GAAGCTGTCA
351 TTGTTAAACA AACTCGGTTA CAATCAACCT TACAACCTTT CGGTTTGTG ACATGAGAAT GCTTTGTCTT
421 TATTCTGTCA GCATGCATTG GGTGAAGATA ACTTCAATTC ACATCCAACA CTTAAACCAC ATGGCGGAGG
491 TATTGTTGAA AAATGTGATG GATTGCCATT GGCAATTGTC ACATGATGAT GATG
```

SEQ ID 137

SEQ ID NO: 3

R1G1C

[Strand]

```
1   TCCCGTGCAA CGTNTATCAT TCAGAAGNGC CCAAAGACCA NAGATNTGTT TAANGNTGNT TMTCAGAAGG
71  AAGTAAATGA TGAAGCTGTN AAAAGATGGC TGATTGATNT CCAACAATTG GCTTAGGACA CTGANGACNA
141 ACTTGATGAT NTCGCCACAG AAGCTATTCA TCGTGAGTTG ATCCGTGAAA CTGGAGCTTC CNGCAGCATG
211 GTAGAAAGC TAATCCCAAG TTGTTGCACA AGTTTCTCAC AAAGTAATAG GATGCATGCC AGGTTAGATG
281 ATATTGCCGC TAAGTNACAA GAACTGGTAG AGGCGAAAAA TAATCTTGGT TTAAGTGTGA TAACATACGA
351 AAAACCCAAA ATTGAAAGAG ATGAGGCGTN TTTGGTAGAT GCAAGTGGTA TCATTGGACG TGAAGATGAT
421 AAGAAAAAAT TGCTTCAGAA GCTGTTGGGG GATACTTATG AATCAAGTAG TCAAACTTC AACATCGTGC
491 CCATAGTTGG TATGGGTGGG GTAGGTAAAA CAACTCTAGC TAGACTTTTG TATGATGAAA AAAAAAGTGA
561 GGATCACTTC GAACTCAGGG TTTGGGTTTG TGTTCCTGAT GAGTTCAGTG TTCCCAATAT AAGCAGAGTT
631 ATCTATCAAT CTGTGACTGG TGAACAACAA GAATTTGCAG ATTTAAATCT GCTTCAAGAA GCCCTTAAAG
701 AGAAACTTCA GAACAACTA TTTCTAATAG TTTTAGATGA TGTATGGTCT GAAAGCTATG GTGATTGGGA
771 GAAATTAGTG GGCCCATTTT ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTACTCG GAAGGAGCAA
841 TTACTCAAAC AGCTGGGTTT TTCTCATGAA GACCTCTGCG ATAGTATAGA CTCCCTGCAA CGTCTATCAC
911 AAGAAGATGC TTTGTCCTTG TTTTCTCAAC ACGCATTTGG TGTACCTAAC TTTGATTTCAC ATCCAACACT
981 AAGGCCATAT GGGGAACAGT TTGTGAAAAA ATGTGGGGGA TTGCCTTTGG CCTTGT
```

SEQ ID NO:4

RLG1D

[Strand]

```
1  CNTACCCATTC TACGAGATCG CTGTCCTCC TCGATCTGCT TAACGATGCT TCCCAGAAGG AAGTNACTAA
71  TGAAGCCGTT AAAAGATGGC TGAATGATCT CCAACATTGG GCTTATGACA TANACGACCT ACTTGATGAT
141 CTTGCAACAS AAAGCTATT CTCSTGAGTT GACCGANGAA GGTGGAGCCT CCACCACTAT GGTAAAGAAA
211 CTAATCCCAA GTTGTTCAC AAGTTTCTCA CAAAGTTATA GGATGCATGC CAAGTTAGAT GATATTGCCA
281 CCAGGTTACA AGAACTGGTA GAGGCAAAAA ATAATCTTGG TTTAAGTGTG ATAACATATG AAAAGCCCCA
351 AATTGAAAGG TATGAGGCAT CTTTGGTAGA CGAAAGTGGT ATTTTGGAC GTTNAGATGA TNAGAAAAAA
421 TTGATGGAGA AGCTGTTGGA GGATAAAGAT GAATCCGGAG TCNAACTTC AGCATCCTGC CCATAATTGG
491 TATGGGTGGA GTTGGCNAAA CAACTCTAGC TAGACTCTTG TTTGATGAAA AGACAGTGAA GGATCACTTC
561 GAACTCAGGG CTTGGGTTTG TGTTCCTGAT GAATTCAGTA TTCTCAACAT AAGCAAAGTT ATCTATCAAT
631 CTGTGACCGG GGAAAGAGAA GAGTTTGAAG ACTTAAATCT GCTTCAAGAA GCTCTTAGAG GGAAACTACA
701 AAACAAACTA TTTCTAATAG TTTTGGATGA TGTATGGTCG GAAAGCTATG GTGATTGGGA GAAATTAGTG
771 GGCCCATTTT ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTIONCG GAAGGAGCAA TTACTIONAAC
841 AGTTGGGTTT TTCTCATCAA GACCTCTGCG GTTGTATAGA CTCCCTGCAA CGTCTATCAC AAGATGATGC
911 TTTGTCTTTG TTTGCTCAAC ACGCATTTGG TGMCCA
```

RLGLE
[Strand]

```
1   TCTAGCTAGA CTTTGTGTATG ACGAGATGCA AGAGAAGGAT CACTTCGAAC TCAAGGCGTG GGTTTGTGTT
71  TCTGATGAGT TTGATATATT CAATATAAGC AAAATTATTT TCCAATCGAT AGGAGGTGGA AACCAAGAAT
141 TTAAGGACTT AAATCTCCTT CAAGTAGCTG TAAAAGAGAA GATTTCAAAG AAACGATTTC TACTTGTTC
211 TGATGATGTT TGGAGTGAAA GCTATGCCGA TTGGGAAATT CTGGAACGCC CATTCTTTC AGGGGCAGCC
281 GGAAGTAAAA TTATCATGAC GACCCGGAAG CAGTCATTGC TAACCAAAC TCGTTACAAG CAACCTTACA
351 ACCTTTCCGT TTTGTACAT GACAGTGCTC TCTCTTTATT CTGTCAGCAT GCATTGGGTG AAGATAACTT
421 CGATTCCAT CCAACACTTA AACCACATGG CGAAGGCATT GTTGAAAAAT GTGCT
```

SEQ ID NO:5

RLG1F
[Strand]

```
1   ATTTTCNGCT  CBAACAAAN  AAAAGCAATG  GCTGAAATCT  TTCTTTGNGC  ATTCTAGACC  AGTATTCTTT
71  GAAAAGNIGG  CTTCTGAAGC  CTTGAAGAAG  ATCGCTCGCT  TCCATCGGAT  TGATTCTGAG  CTCAAGAAAC
141 TGAAGAGGTC  ATTAATCCAG  ATCAGATCTG  TGCTTAATGA  TGCTTCTGAG  AAGGAAATAA  GTGATGAAGC
211 TGTAAAGAA  TGGCTGAATG  GTCTCCAACA  TTTGTCTTAC  GACATAGACG  ACCTACTTGA  TGATTTGGCA
281 ACCGAAACTA  TGCATCGTGA  GTTGACCCAC  GGATCTGGAG  CCTCCACCAG  CTTGTAAGAA  AGATAATCCC
351 AACTTGTTC  ACAGATTTCT  CACTAAGTAG  TAAGATCCGT  AACCAAGTTAG  ATAATATTAC  CATCAAGTTA
421 CAAGAAGTGG  TAGAGGAAAA  AGATAATCTT  GGCTTAAGTG  TGAAGGTTGA  AAGCCCAAAA  CATACCAACA
491 GAAGATTACA  GACCTCTTTG  GTAGATGCAT  CTAGCATTAT  TGGTCTGTGA  GGTGATAAGG  ATGCATTGCT
561 CCATAAGCTG  CTGGAGGATG  AACCAGTGA  TAGAAACTTT  AGCATCGTGC  CAATAGTTGG  TATGGGTGGT
631 GTGGGTAAGA  CGACTCTAGC  TAGACTTTTG  TATGACGAGA  TGCAAGAGAA  GGATCACTTC  GAACTCAAGG
701 CGTGGGTTTG  TGTTCCTGAT  GAGTTTGATA  TCTTCAATAT  AAGCAAAGTT  ATCTTCCAAT  CGATAGGTGG
771 TGGARACCA  GAATTTAAGG  ACTTAAATCT  CCTTCAAGTA  GCTGTAAAAG  AGAAGATTTC  AAAGAAACGA
841 TTTCTTWTG  TTCTGGATGA  TGTTTGGAGT  GAAAGCTATA  CAGAATGGGA  AATTCTAGCA  CGTCCATTTC
911 TTGCAGGGGC  ACCAGGAAGT  AAGATTATCA  TGACGACCCG  GAAGTTGTCG  TTGCTAACCA  AACTCGGTTA
981 CAATCAACCT  TACAACCTTT  CSGTTTGTGC  ACATGATAAT  GCTVTGTCIT  TATTCTGTCA  GCAYGCATTG
1051 GGTGAAGATA  ACTTCGATTC  ACATCCAACA  CTTAAACCAC  ASGGTGAAAG  TATTGTTGAA  AAATGTGACG
1121 GTTTACCATT  GGCTTTRATT  GCACTTGGGA  GRTTGTGAR  GACAAAAACA  GATGAGGAAG  AATGGAARGA
1191 AGTGTGTAAT  AGTGAAATAT  GGGGGTCAGG  AAAGGGAGAT  GAGATTGTC  CGGCTCTTAA  ACTAAGCTAC
1261 AATGATCTCT  CTGCTCTTT  GAAGAAGTTG  TTTGCATACT  GCTCCTTGTT  CCCAAAAGAC  TATGTGTTCG
1331 ATAAGGAGGA  GTTGATTTTG  TGTGTGATGG  CAGAAGGGTT  TTTGCACCAA  TCAACCACAA  GCAAGTCBAT
1401 GGAACGCTTG  GGHGATGAAG  GTTTTGATGA  ATTGTTGTCA  AGATCATTTT  TTCAACATGC  CCCTGATGCC
1471 AAATCGATGT  TTGTGATGCA  TGACCTGATG  AATGACTTGG  CHACATCTCT  TGCTGGAGAT  TTTTMTTCAA
1541 GGATGGACAT  TGAGATGAAG  AARGAATTTA  GGAAGGAAGC  TTTGSAAAAG  YAYCGCCATA  TGTCAMTTGT
1611 TTGTGAKGAT  TACATGGTK  ACAAAGGTT  CRAGCCATTS  ACAAGGAGCT  AG
```

SEQ ID NO: 6

RLGIG
[Strand]

```
1  GTGAAGGATC ACTTCGAACT CAGGGCTTGG GTTTGTGTTT CTGATGAATT TAATATCCTC AATATAAGCA
71  AAGTAATTTA TCAATCTGTA ACCGGGGAAA AAAAGGAGTT TGAAGACTTA AATCTGCTTC AAGAAGCTCT
141 TAAAGAAAAA CTTTGGAATC AGTTATTTCT AATAGTTCTG GATGATGTGT GGTCTGAAAG CTATCGTGAT
211 TGGGAGAAAT TAGTGGGCCC ATTTTITTCG GGGTCTCCTG GAAGTATGAT TATCATGACA ACTCGGAAGG
281 AGCAATTGCC AAGAAAGCTG GGTITTCCTC ATCAAGACCC TTTGCAAGGT CTATCACATG ACGATGCTTT
351 GTCTTTGTTT GCTCAACACG CATTTGGTGT ACCA
```

SEQ ID NO: 7

RLGH
[Strand]

```
1  TCTAGCTAGA CTTTGTATG AGGAAATGCA AGGGAAGGAT CACTTCGAAC TCAAGGCGTG GGTATGTGTT
71  TCTGATGAGT TTGATATCTT CAATATAAGC AAAATTATCT TACAATCGAT AGGTGGTGGA AACCAAGAAT
141 TTACGGACTT AAACCTGCTT CAAGTAGCTT TAAAAGAGAA GATCTCAAAG AAAAGATTTC TTCTTGTTC
211 TGATGATGTT TGGAGTGAAA GCTATACCGA TTGGGAAATT CTAGAAGGCC CATTTCCTGC AGGGGCACCT
281 GGAAGTAAGA TTATTATCAC CACCCGGAAG CTGTCATTGT TAAACAACT CGGTTACAAT CAACCTTACA
351 ACCTTTCGGT TTGTGCACAT GAGAATGCTT TGTCTTTATT CTGTCAGCAT GCATTGGGTG AAGATAACTT
421 CAATTCACAT CCAACACTTA AACCACATGG CGAAGGTATT GTTGAAAAAT GTGAT
```

SEQ ID NO: 8

RLGI
[Strand]

```
1  TCTAGCTAGA CTTGTGTATG ATGAGATGCA AGAGAAGGAT CACTTTGAAC TCAAGGCGTG GGTATGTGTT
71  TCTGATGAGT TTGATATATT CAATATAAGC AAAATTATTT TCCAATCGAT AGGAGGTGGA AACCAAGAAT
141 TTAAGGACTT AAACCTCCTT CAAGTAGCTG TAAAAGAGAA GATTTTAAAG AAACGATTTC TTCTTGTTC
211 TGACGACGTT TGGAGTGAAA GCTATGCCGA TTGGGAAATT NTGGAACGCC CATTTCTTGC AGGGGCAGCC
281 GGAAGTAAAA TTATCATGAC AACCCGAAAG CAGTCATTGC TAACCAAACT CGGTTACAAG CAACCTTACA
351 ACCCTTCCGT TTGTACACAT GACAGTGCTC TGTCTTTATT CTGTCAGCAT GCATTGGGTG AAGGTAACCT
421 CGATTACACAT CCAACACTTA AACCACATGG CGAAGGCATT GTTGAAAAAT GTGCTGGATT GCCATTGGCA
491 TTGTCGACA
```

SEQ ID NO. 9

RLGLJ

[Strand]

```
1   TACTACTACT AGAATTCGGT GTTGGTAAGA CGACTCTAGC TAGACTTTTG TATGAGGAAA TGCAAGGGAA
71  GGATCACTTC GAACTTAAGG CGTGGGTATG TGTTTCTGAT GAGTTTGATA TCTTCAATAT AAGCAAAATT
141 ATCTTACRAAT CGATAGGTGG TGGAAACCAA GAATTTACGG ACTTAAACCT GCTTCGAGTA GCTTTAAAAG
211 AGAAGATGTC AAAGAAAAGa TTTCTTCTTG TTCTTGATGA TGTTTGGAGT GAAAGCTATA CCGATTGGGA
281 AATTINTAGAA CGCCCATTTT TTGCAGGGGC ACCTGGAAGT AAGATTATTA TCACCACCCG GAAGCTGTCA
351 TTGTTAAACA AACTCGGTTA CAATCAACCT TACAACCTTT CGGTTTGTG ACATGAGAAT GCTTTGTCTT
421 TATTCTGTCA GCATGCATTG GGTGAAGATA ACTTCAATTC ACATCCAACA CTAAACCAC ATGGCGGAGG
491 TATTGTGAA AAATGTGATG GATTGCCATT GGCATTGTG ACATGATGAT GATG
```

SEQ ID NO:10

RLGIA a.a.

IVTVRTR?LSLLHLLSYVIFS?I?PH?ILWLF.INFYSTCHFMSFSILLSFT.YLNVITINAYLFFFK.THIYR
LKSynt.VKLI.YICSSFVYLYVSSLIYLLFIY.SR.SL.Y.KFNLFKI.NY..SHNLNKIKKNGPTISPSLFQLINIV
SILLRFHPNQYFQRMDSYGVSEFAFRHCSLKEIINQMELLQCSLLMKGELYVK?MSAI?LHPEPTTSLV
YHQTQNGGSR?T?KS.RIDYFCPHGLTEERV.FIIFL?KNYRSIEFLHRQRSFTSQCFVKQFLIFLSFR.NS
SIATCNLQLLGPQICGGR.FNPHIHCKQ.FKSISVHPIHQHLLIIEIIHASSISSTSILYSLLLSY.TMAEIVLS
AFLTUVFEKLA?EALKKIVRSKRIESELKKLKETLDQIQDILLNDASQKEVTNEAVKRWLNDLQHLAYDID
DLLDD?ATEAV?RELTEEGGASSSMVRKLIPSCCTSFSQSNRMHAKLDDIATRLQELVEAKNNLGLSVI
TYEKPKIERYEASLVDESGTVGREDDKKKLEKLLGDKDESGSQNFSIVPIVGMGGVGKTTLARLLYDEK
KVKDHFELRAWVCVSEFSVPNISRVYQSVTGEKKEFEDNLILQEALKEKLRNQLFLIVLDDVWSESY
GDWEKLVGPFLAGSFGSRIIMTTRKEQLLRKLGFSHQDPLEGLSQDDALSIFAQHAFGVPNFDShPTLR
PHGELFVKKCDGLPLALRTLGRLLRTKTDEEQWKELLDSEIWRGKSDIWPALRLSYNDLSA?LKLLFA
YCSLFPKDYEFDKELLLWMAEGFLHQPT?NKSQRGLLEYF?ELLSRSFFQHAPN?KSLFVMHDLMDND
LATFVAGEFFSRLDIEMKKEFRM?SLEKHRHMSFVCE?YIGYK?FEPFRGAKNLRTFLALSVGVVEDWK
MFYLSNKVLND?LQDLP LLRVL?LI?L?I?VP??VGSM?HLRYLNLS?T?ITHLPE??CNLYNLQTLIV
SGC?YLV?LPKTF?LKNL?HFDMR?TP?LKNMPL?IGELK?LQTLF?NIGIAITELKNL?NLHGK?CIGG
LGKMENAVGCTLSELVSKKV?..?NW??G...I.CFPKWEHLKKKSSMK.CLIMVL?KKP?IMSIGGIEFPN
WVGLRVSETRDVFMVYEK?CFT.FHQSPSGK.MIFSG?TDEMWRGMIG?LGAVEEISIHSCNEIRYLWE
SEAEASKVLMNLKKLDLGECENLVSLGEKKEDNHNINSGSSLTSFRRLNWWRCNSLEHCRCPDSMENLY
MHMCDS?TSVSFPTGGGQKIKSLTITDCKKLSEELGGRETRV LINSKMQMLESVDIRNWP NLKSISEL
SCFIHLNRLYISNCPS?ESFPDHELPNLTSLTDRRRGQRFSYERLRFDWPSF

SEQ ID NO:11

RLGIB a.a.

NRSYENRCPLLPVI...EKI.LKV.IQNPLFHR.YDALAVVCKNSIINANQNSS.ETI.IMV.IVLSPTYTHFFQIPII
HTYKCSHIRFSLAMAEILGSAFFAVFFFEKLASEALKRVACSKVIDKELEKLNSS.INIKALLNDASQKEIS
KEAVKEWLNALQHLPYDIDDLGDLATKAIHRKFSEEGATINKVRKLIPSCFSSLSSTKMRNKIHNITS
KLQELLEERNNLGLCEIGESRKLNRKSETS?LDPSSIVGRTDDKEALLKLYEPCDRNFSILPIVGMGGL
DKTTLGRLLYD?MQVKDHFELKAWVCVSDEFDIFGISKTFESIEGGNQEFKDLNLLQVALKEKISKKRFL
VLDDVWSESYTDWEILERPFLAGAPGSKVIITTRKLSLLNQLGHDQPYQLSDLSHDNALSIFCQHAFG
VNSFDSHPILKPHGEGIVEKCDGLPLALIALGRLLRTKRDEEWEKELLNSEIWRGKRDEIIP?LRLSYND
LSASLKQLFAYCSLFPKDYVFNKEKLILLWMAEGFLHNENTNKSMEERL?LEYFDDLLSRFFQHALDDKS
LFVVHDLMDLATS VAGDYFLRLDIEMKKEALEKYRHMSFVCESYMYKRFEPFKGAKKLRTFLAMPV
GMIKSWTTTFYLSNKVLDLLHELPLLRVLSLSYLSIKEVPEIIGNLKHLRYLNLSHTSITHLPENVCNLYN
LQTLILCGCCFITKFPNNFLKLRNLRHLDISDTPGLKKMSSGIGELKNLHTLSKLIIGGENRLNELKNLQNL
H

SEQ ID NO:12

RLG 1 c.a.

SRAT?IIQK?PKT?D?F????QKEVIDEAVKRWLID?QQLAYDT?D?LDD?ATEAIHRELIRETGAS?S
MVRKLIPSCCTSFSQSNRMHARLDDIAAK?QELVEAKNNLGLSVITYEKP KIERDEA?LVDASGIIIGRED
DKKKLLQKLLGDTYESSQNFNIVPIVGMGGVGKTTLARLLYDEKKVKDHFELRVWVCVSDEFSVPNIS
RVYQSVTGENKEFADLNLLQEALKEKLQNKLFLLVDDVWSESYGDWEKLVGPFHAGTSGSRIIMTTR
KEQLLKQLGFSHEDPLHSIDSLQRLSQEDALSLFSQHAFGVPNFDSHPTLRPYGEQFVKKCGGLPLAL

SEQ ID NO:13

RLG ID

?T?LRDRCPSSICLTMLPRRK?LMKPLKDG.MISNIWLMT?TTYLMILQ?KAI??ELT?EGGASTSMVRK
LIPSCCTSFSSQSYRMHAKLDDIATRLQELVEAKNNLGLSVITYEKPPIERYEASLVDESGIFGR?DD?KK
LMEKLLEDKDESGVKLQHLPIIGMGGVG?TTLARLLFDEKTVKDHFEFRAWVCVSDEFSILNISKVIYQS
VTGEKKEFEDLNLLQEALRGKLQNKFLIVLDDVWSESYGDWEKLVGPFHAGTSGSRIIMTTRKEQLLK
QLGFSHQDPLRCIDSLQRLSQDDALSFAQHAFG?

SEQ ID NO: 14

RLGIE

LARLLYDEMQEKDHFELKAWVCVSDEFDIFNISKIIFQSIGGGNQEFKDLNLLQVAVKEKISKKRFLLVLD
DWSESYADWEILERPFLAGAAGSKIIMTTRKQSLTLKGYKQPYNLSVLSHDSALSFLCQHALGEDNF
DSHPTLKPHGEGIVEKCA

SEQ ID NO: 15

RLGIF

FSA?NK?KQWLKSFF?HSRPVFFEK?ASEALKKIARFHRIDSELKKLKRSLIQIRSVLNDASEKEISDEA
VKEWLNGLOHLSYDIDDLDDLATETMHRELTDDLEPPPACKKDNPTCCTDFSLSSKMRNKLDNITIKL
QELVEEKDNLGLSVKGESPKHTNRRLQTSYDASSIIGREGDKDALLHKLLEDEPSDRNFSIVPIVGMGG
VGKTTLARLLYDEMQEKFELKAWVCVSDEFDIFNISKVIFQSIGGG?QEFKDLNLLQVAVKEKISKKR
FL?VLDDVWSESYTEWEILARPFLAGAPGSKIIMTTRKLSLLTKLGYNQPYNLSVLSHDNALSFCQHA
LGEDNFD SHPTLKP?GESIVEKODGLPLALIALGRLL?TKTDEEEWKEVLNSEIWGSGKGD EIVPALKLS
YNDLSASLKKLFAYCSLFPKDYVFDKEELILLWMAEGFLHQSTTSKSMERLGHEGFDELLSRFFQHAPD
AKSMFVMHDLMDLATS VAGDFFSRMDIEMKKEFRKEAL?K?RHMS?VC?DYMV?KRF?P?TRS.

SEQ ID NO: 16

RLG1 G

VKDHFELRAWVCVSDEFNILNISKVIYQSVTGEKKEFEDLNLLQEALKEKLWNQLFLIVLDDVWSESYR
DWEKLVGPFFSGSPGSMIIMTTRKEQLPRKLGFPHQDPLQGLSHDDALSLFAQHAFGVP

SEQ ID NO: 17

RLG 1 H

LARLLYEEMQGKDHFKAWVCVSDEFDIFNISKIILQSIGGGNQEFTDLNLLQVALKEKISKKRFLLVLD
DVWSESYTDWEILERPFLAGAPGSKIITTRKLSLLNKLGYNQPYNLSVLSHENALSLFCQHALGEDNFN
SHPTLKPHGEGIVEKCD

SEQ ID NO: 18

RLGI I

LARLVYDEMQEKDHFELKAWVCVSDEFDIFNISKIIFQSIGGGNQEFKDLNLLQVAVKEKILKKRFLLVLD
DVWSESYADWEI?ERPFLAGAAGSKIIMTTRKQSLTKLGYKQPYNLSVLSHDSALSFCQHALGEGNF
DSHPTLKPHEGIVEKCAGLPALST

SEQ ID NO: 19

RLG 15

EFGVGKTTLARLLYEEMQGKDHFELKAWVCVSDEFDIFNISKIILQSIGGGNQEFTDLNLLRVALKEKISK
KRFLVLDDVWSESYTDWEI?ERPFLAGAPGSKIITTRKLSLLNKLGYNQPYNLSVLSHENALSIFCQH
ALGEDNFNSHPTLKPHG?GIVEKCDGLPLALS

SEQ ID NO: 20

SEQ ID NO: 21
RLG 2A

```
1   TTNACACCAT AAATTCCTCNA CCTGNGGGGA CAAAAACCTA AAAATGGTCC ATAATGCNCA AATCAGNAAG
71  GTTGANAAG  CTCTAAGTTT  TTNACCTCCA NCTGATGCNC NNTCCTCNIA AAGTTCANAT CCAAGCTTGC
141 CCTCCAACCTC TANCNCCTTC AATGGCACCT CCTTCTCTTC AAAAGCACAC AAGAACACIT TCAAGCTCAA
211 CCACACTCAC  ACAAGCTCTA GAACNAGGGT TAGGGCACAT TTAGGGTTTT GCTCTCTGGA AATGGTGTCT
281 AAAAGTGAGG  CCATAATGTT CCTTATATAA GGCTCACTCC CACAATTAGG CTTTCAATCT GAACGTANTA
351 CGCCCACTGT  ACACATAGGT ACGCCCAACG TACTCGGTAG TCTCCGCGTC AANAATACAC TCATGAGTAC
421 CGGCAACGTA  CTTTCCCTTA CGCCCAAGCGT ACTCAAAGC  CAAACATTCT TTTCAAGGAC TAATTTTGAC
491 AACTTGAAGGA AAGAAAAGGA TCAAAGANAT ATACTTGAAT TCCGGGATGT TACAATGAAG TTGANACCTT
561 GGCTAAAAAA  TTAATTTGGT TGTGGAAGCC GTTGGCTGAG CAAGCAACAA GGGTAAAAAT CGTAATCTAC
631 AAATGGTGT  ATTTCCTATT TCTTCTTATT ATTTTACTTG ATTTACGGGT AGTTTTTTTT TCTTACAAAA
701 AATATTAAG  TTGATAAAGT ATAGCCACTA AAATTGACTT TTTCCAAAAC ATAATGTCAA ATGGTGCCTA
771 TATGTATCAT  GTTGATTAN  ATAATGAATA TGATGATNCT GTTCTATTTA ANCCGAAAAA ATTATCTAAT
841 GATTTTATAT  TGGAAAACAA AGTTGTGATT TTTNGCATAA TATAATCAA  TCCNCTTTTG TNGGGAGGT
911 GGATAAATGT  GGTAAATTTA NAACAAGTGT TTNACNITG AAGGGTWTGG AAAGGTGAA  AAAAGTTAA
981 ATGATAAAAT  GTTTACACAA ATGTTGTATC CGACTGAATA TNAITTTAA  GGATNATTGT ATTAATTTGT
1051 TGATATATAG  TAAGCATAAA TATTTAGAAT TGTGACTTAA ATTTATAAGT TATNCNACT  GGATTGAAAC
1121 ATTTTGTATA  TANATTAGGA ATGAAAATGA GCAACCCATA CATACTTATC TTTGGTAGIT TGGTTATTAT
1191 ATTTTATTA  NAATATAGAA NCATCCCTTT ATTTTAAACC CATATGTGG  ACGGACTTGA ATAAATGGGA
1261 AAAATGTACC  TTGCTATTTA GCACAAAAAA ATTATAAAAA TGTACATTGC TATTTAGCAC AAACAAAAAA
1331 AAAAAACTTA  TCCTTTTGTG ATTAGGTAC  AAAGAAATAT AAAATGGGAA ATGTGTGTCT ATTTAATGCA
1401 CTAAAAGAAA  CTATTTTGCC TTTATTAAAC CGGGTAAACC AATAGAAAA  TGAAGTACA  TTGTCATTTA
1471 GCATGAAAA  AAATAACTTT CCATTTTTTG CATCCGGTCA CAATAATAGA AAAATGAAAG TACGTGTCTA
1541 TTTAGCGAAA  CTAACCTTCT TTTTCTTTTT TGGCATCGTA TCATAAAATA TAGACTAAA  TACGTTAGTT
1611 TTACATTTTT  AATACATTGA AATGCTAAT  CCACATGTTA TTCTATAAAA AGGGAAATGT AATTTACTTA
1681 TTCTTTGATT  CTTTGGCTTC TTTTGTAGTAC CCAAAACATC CCTCTATCCA TCTATTTCAA CTAATAAAT
1751 GAAAACTATA  TTCTTCCAT  TGTAGGGATG TTATAAATTT TGTAAATGTT TTTATGCAAA AAAGTGTITT
1821 TIGTTAACTA  GATTAACGAG ATTCATTTTT CAGCATTTTA GGAGAAGTTC ATCCATCTTT TGGATATGAA
1891 GTGCAAGCCA  AGTTCTTTAA CATGGAATAT GAGGTCCCTA TATGCTCAAA AAATAGCAAA TGAAGAAAT
1961 TTTAAATGG  ATCCCATATA AAGAAAAATTT GTTAATGGTT GTTTTAATAT TGGTCAATGT GTCCACCGGA
2031 TGAGCATTA  ACTAGTTTAT AAGGGTAAA  GGTGGGTTTG GTGGGCCCAT TTATCTTTAT TATTTCTAAA
2101 AGTCAGAATT  AAGTAAAAAA AATTATAAGA TAAATACCAT AAGGATAAAA AATCATTTTA TTTGGACCAA
2171 AGACCAAAG  TGTAAAGGGG CTGTTTGTIT TTTTGTGAA  GAGCTGTGCA ACCACTTTTG TCTGCGCCGC
2241 ACAGACAAG  TGCAGACATA TGCCCTCGCA GAGTGTTTGT TTTTGTAAAG TGCGCAGACC AAAAAACGT
2311 CTGCGCGAG  TGCCTCTGGC GCATATATGT GTCACGTGCT TCAAAGGTCT TCAGACCTCA TTTTAACCAA
2381 AAAAAAATA  GACCACCGGT TTTTTTTTTT TTTTNTTCT  TTCTCTTGT  GCIGAAAATG CATTTTTAAT
2451 CTTTATGACA  TGAATTAAG  TTTGAAAAAT TAATTTATTT CAACAGCTGT AGACGTAAA  AACAAACAGT
2521 CTTCCTGTT  CAGACTGTGG ACCTCTTCT  ACCGCAGAGA CTTGCAGAT  TGGTCCGAG  TGGTCCGAG
2591 ACTGCAAGCA  TTTTGGCTTC AAATAACAA  ACATCACCTA ATTTGACTAC ACCACACGGA CCTCCAATGT
2661 AACAAAAAA  AGGTTGAAC  AAAGTTGCC  ATTTCTCCAT ATCCAGGGGC CATTTATGTA AGAGTTATCT
2731 AAATTTTAGT  TCGGTAGATC AGTTCTCACA TTTTAACCGG GTAAAGTGT  TGTGTGTACG CGCGCACCTG
2801 AAAGGTTTGA  ANGTAAC TTC  CAACTGAAN  CAANAATCGA TATGAAGTAT CAAGTTAGAG GTTCAATTGG
2871 TGAAGGAATC  AGCTGGAGGT TGGGGAATCG AGCTTCCACT ATTAAGGTAA AATCCATAAC CCTAAATGTT
2941 GGTACGCTCA  TATATCAAAT TGCCTGTTTT GTTGAATGAA AAAAGCATGC TCAAAAAACC AGTGTAAAGC
3011 ACGGTATATG  ACATATTTAT AGTTACTGAT AACAAATTAT GATAATTTTG GGTTTACGTA AGTTAGGATT
3081 CGTACTTCAA  CCAAAATGTA TAGTTTTTGT GAGTCTATCT ATGTATTTGG GGAATCACAT TAGCAACGGG
3151 ATGTACTAG  TAATTCGAAA AAGTCTTTTA AATAATTTT  CTGTTTATA  TTTATGAATA GTTTTAGCGA
3221 CATCTAATAT  TAAATAGAAT GTATCTGATA TTGAATTAAT GTCCTTAATG TGAACATAGA CCTTTTCCAT
3291 TTTACTAATGC  CTAATTATTA GTTTCTAATC AATAAATTTT AATTTCTGTT TTAGCTTCT  AAGACAATAA
3361 AAATCCATGA  TTTACCTTTA AATATTAA  AAAATGACCA TAAATRAATA AAAAATTAGG ATACCAAACC
3431 CCCCCGCCAT  GCCCAATGTC TAAATATTCT TGATGCTTTT GCTTTTCCCT CTTTCTCTTG TTAGTCTATT
3501 ATTCTGAGGA  GTTTGAGAGA GTTTCATACA AGAAAATTT  AAGAAGAAAG CAAAGGTCCA GGTATTTCTCT
3571 TTTCTTAATT  ATGTATTAA  TTACAAGCAT TTTTACACG  ATCCATGGTT TTTTGTGTAT GTTTTTCAAA
3641 TTGAAACTAG  ATTTGGACTT TTGCCCTTGA TGATTCATAA GATATTGCAT GGAGTTGAGA TTGTGTGAAG
3711 AAAGTGGTGA  ATAGAAAGAG CAAGTGAATC CAGATATAGT ATTTGTAATA TATGATGATG AGATAGAGAT
3781 ATGTTAAAC  TGGCTAGAAA ATTTGTTTAA TTTGAAATTT AGGTGTGTGA ATTTGAAAGA TACCAAGCTA
3851 ATAACTAATT  AGTATGCTA  AATAGTTATA AAGAACAACA AACTCGTAGT TTTTCTTCA  TGAATTTCAA
3921 CCTCTTCTGA  CCAAACTAAA TTATAACAAA ATTTGAATATC ATTTCTCTGA ATCAATTTTA ACTTTTGTTA
3991 TTATCAATCAT  GTCTAAAATT GCCACAAGTT TATTTTCATA GTCATATTGG ATTATGAAAG GACTATTTTT
4061 ACCAATTACA  TCTTTACTTT ATGGCCAAAG CTAATACAAT CCGACTAAAC TAAAGGATTC TAGGATGCAT
```


SEQ ID NO: 2/
RLG 2A cont.

```
4131 ATAGTTTGGCT CCCGATTAT AGATTTCAT CTAATTGTG TATTGTACTA ATTTAGGTGC CACCACAAGT
4201 AAATTCCTGA AATGGATGTC GTTAATGCCA TTCTTAAACC AGTTGTGAG ACTCTCATGG TACCCGTTAA
4271 GAAACACATA GGGTACCTCA TTTCCTGCAG GCAATATATG AGGGAATGG GTATCAAAAT GAGGGGATTG
4341 AATGCTACAA GACTTGGTGT CGAAGAGCAC GTGAACCGGA ACATAAGCAA CCAGCTTGAG GTTCCAGCCC
4411 AAGTCAAGGG TTGGTTTGAA GAAGTAGGAA AGATCAATGC AAAAGTGGAA AATTTCCCTA GCGATGTGTG
4481 CAGTTGTTTC AATCTTAAGG TTAGACACGG GGTCCGAAAG AGAGCCTCCA AGATAATTGA GGACATCGAC
4551 AGTGTATGA GAGAACAATC TATCATCATT TGGAAATGATC ATTCCATTCC TTTAGGAAGA ATTGATTTCA
4621 CGAAAGCATC CACCTCAATA CCATCAACCG ATCATCATGA TGAGTTCCAG TCAAGAGAGC AAACCTTTCAC
4691 AGAAGCACTA AACGCACTCG ATCCTAACCA CAAATCCAC ATGATAGCCT TATGGGGAAT GGGCGGAGTG
4761 GGAAGAGACA CAATGATGCA TCGGCTCAAA AAGGTTGTGA AAGAAAAGAA AATGTTTAAT TTTTATTAATG
4831 AGGCGGTGT AGGGGAAAAA ACAGACCCCA TTGCTATTCA ATCAGCTGTA GCAGATTACC TAGGTATAGA
4901 GCTCAATGAA AAAACTAAAC CAGCAAGAAC TGAGAAGCTT CGGAAATGGT TTGTGGACAA TTCTGGTGGT
4971 AAGAAGATCC TAGTCATACT CGACGATGTA TGGCAGTTTG TGGATCTGAA TGATATTGGT TTAAGTCCTT
5041 TACCAATCA AGGTGTCGAC TTCAGGTGT TGTGTGACAT ACGAGACAAA GATGTTTGCA CTGAGATGGG
5111 AGCTGAAGTT AATCAACTT TTAATGTGAA AATGTTAATA GAAACAGAAG CACAAAGTTT ATTCCACCAA
5181 TTTATAGAAA TTTCCGATGA TGTGTATCCT GAGCTCCATA ATATAGGAGT GAATATTGTA AGGAAGTGTG
5251 GGGGTCTTACC CATTGCCATA AAAACCATGG CGTGTACTCT TAGAGGAAAA AGCAAGGATG CATGGAAGAA
5321 TGCACTTCTT CGTTTAGAGC ACTATGACAT TGAAAAATAT GTTAATGGAG TTTTAAATAT GAGTTACGAC
5391 AATCTCCAAG ATGAGGAGAC TAAATCCACC TTTTGTCTT GTGGAATGTA TCCCGAARAC TTTGATATTG
5461 TTACCGAGGA GTTGGTGAGG TATGGATGGG GGTGAAATTT ATTTAAAAAA NIGTATACTA TAGGAGAAGC
5531 AAGAACCAGG CTCACACAT GCATTGAGCG GCTCATTCAT ACAAATTTGT TGATGGAAGT TGATGATGTT
5601 AGGTGCATCA AGATGCATGA TCTTGTTCGT GCTTTTGTIT TGGATATGTA TTCTAAAGTC GAGCATGCTT
5671 CCATTGTCAA ACATAGTAAT ACATAGATAG GGCATGCAGA TAATATGCAC GACTCTTGTA AAAGACTTTC
5741 ATTAACATGA AAGGTAATGT CTAAGTTTCC TACAGACCTG AAGTTTCCAA ACCTCTCCAT TTTGAAACTT
5811 ATGCAATGAA ATATATCAAT GAGGTTTCCC AAAAATTTTT ATGAAGAAAT GGAGAAGCTT GAGGTTATAT
5881 CCTATGATAA AATGAAATAT CCATTGTCTC CCTCATCACC TCAATGTTCC GTCAACCTTC GCGTGTITCA
5951 TCTACATAAA TGCTCGTTAG TGATGTTTGA CTGCTCTGTG ATTGGAATC TGTCGAATCT AGAAGTGTCT
6021 AGCTTGTCTG ATTCGCCAT TGACCGGTTG CCTTCCACAA TCGGAAAGTT GAAGAAGCTA AGGCTACTGG
6091 ATTTGACGGA TTGTTATGGT GTTCGTATAG ATAAATGGTGT CTTAAAAAAA TTGGTCAAACT TGGAGGAGCT
6161 CTATATGACA GTGGTTGATC GAGGTGCAAA GCGGATTAGC CTCACAGATG ATAACCTCAA GGAGATGGCA
6231 GAGCGTTCAA AAGATATTTA TGCAATAGAA CTTGAGTTCT TTGAAAACGA TGCTCAACCA AAGAATATGT
6301 CATTTGAGAA GCTACAACGA TTCCAGATCT CAGTGGGCG CTATTTATAT GGAGATTCCA TAAAGAGTAG
6371 GCACTCGTAT GAAAACACAT TGAAGTTGGT TCTTGAAAAA GGTGAATTTT TGAAGCTCG AATGAACGAG
6441 TTGTTTAAAG AAAACAGAGG GTTATGTTTA ATGTGTTGAG ATATGAATGA TCTTGAAGAT ATTGAGGTTA
6511 AGTCACTCTC ACAACTTCTT CAATCTTCTT CGTTCAACAA TTTAAGAGTC CTGTCTGTTT CAAAGTGTGC
6581 AGAGTTGAAA CACTTCTTCA CACCTGGTGT TGCAAACTT TAAAAAAGC TTGAGCATCT TGAAGTTTAC
6651 AAATGTGATA AATGGAAGA ACTCATACGT AGCAGGGGTA GTGAAGAAGA GACGATTACA TTCCCAAGC
6721 TGAAGTTTAT ATCTTGTGT GGGCTACCAA AGCTATCGGG TTTGTGCGAT AATGTCAAAA TAATTGAGCT
6791 ACCACAATC ATGGAGTTGG AACTTGACGA CATTCAGGT TTCACAAGCA TATATCCCAT GAAAAAGTTT
6861 GAAACATTTA GTTTGTTGAA GGAAGAGGTA AATATAAAT TTTAATGCTA ATACATTACA AAGGATCTTT
6931 TCAGTTAAAT CTTTCAAAAT ATATTGTAAT TTGATTGTAT GGGGTATTAT TGTGGATGG GACTATTAAAT
7001 AAATGATTAT CTGTCAGGTT CTGATTCCCTA AGTTAGAGAA ACTGCATGTT AGTAGTATGT GGAATCTGAA
7071 GGAGATATGG CCTTGCAGAT TTAATATGAG TGAGGAAGTT AAGTTCAGAG AGATTAAAGT GAGTAAGCTG
7141 GATAAGCTTG TGAATTTGTT TCCGCACAAG CCCATATCTC TGCTGCATCA TCTTGAAGAG CTTAAAGTCA
7211 AGAATTTGTTG TTCCATTGAA TCGTTATTCA ACATCCATTT GGAATGTGTT GGTGCAACTG GAGATGAATA
7281 CAACAACAGT GGTGTAAGAA TTATTAAAGT GATCAGTTGT GATAAGCTTG TGAATCTCTT TCCACACAAT
7351 CCCATGTCTA TACTGCATCA TCTTGAAGAG CTTGAAGTCG AGAATTTGTTG TTCCATTGAA TCGTTATTCA
7421 ACATTGACTT GGATGTGCTT GGTGCAATTG GGCAAGAAGA CAACAGCATC AGCTTAAGAA ACATCAAAGT
7491 GGAGAATTTA GGAAGCTAA GANAGGTGTG GAGGATAAAA GGTGGAGATA ACTCTCGTCC CCTTGTTCAT
7561 GGCCTTCTCAAT CTGTTGAAAG CATAAGGTTT ACNAAATGTN AGAAGTTTAG AAATGTATTC AACCTTACCA
7631 CCACAATTTT TAATCTGGGG GCACTTTTGG AGATTTCAT ATAGTACTGC GGAGAAAACA GGGGAATGA
7701 CGAATCGGAA GAGAGTAGCC ATGAGCAAGA CGAGGTAAGG ATTTCAATTT CACTGTCTTA ATTAATGATT
7771 AAGCTCTCTG TTTTGAATA AAAAAGGCAC AAACCATTTT ATGACTTAAT GTAGCAATAC AAGTCATGTA
7841 TAAGAGTGAC CAACTCTTTT TTATTTATAA AATGACTACA AAATATTTTT TTTTATTAGA GATCATGTAT
7911 AAATGTGACT AATTTTTCAT CACCTAACTT TAGTTGATAA ATCTTTATAA ATGTCACTAG TTACTTTTCA
7981 GTAAAATGAC AAATTTAATA AATTATCAAC AAAAAGCATC AACTCAAAAA ATCCCAACAC CCGTAAATAT
8051 TTAATAATAA AGGATTTAAC ATCTAATACC AACAATTTTT TTTCTAAACA TGATTTGGAC CAAATATCAC
8121 CAGCAACTCA AGTTTGGAAAT CGATTACGCT TAAACTTGA CCAGCATAAT TAGATAGATG AGAGTTGAAG
8191 CTAAAGTGCC TATATAAGTT CGTTTCATCT TTTTCTTGA TCTTGATAGC AAGTTGAATG ATTTTCTTCT
```

RLG2A cont.

8261 TCAAAATTGA TAAAAATCTA CATTATAAAG AGACTAGCTT GAAAAAAAT GGTCTAGGTG GGTCTTGGGT
8331 TCTGGTAGAT GAAGATGGAA GGGGAGAGTA TGATTTCAAA GACACAACAC ATCCTTCATT TTATTTATTT
8401 ATTATTTATTA TTATTTTGTG ATATCTTGCT CATATTTGTT ACAGATATGT GAGGTCTATT AATCTTTTGA
8471 AATATATAAA AAAATAAATA ACATAAATGA GAAAATTAAA TAAAGAATAA ATTAATAAGG GCACAATAGT
8541 CTTTTTAGGT AAGACAAGGA CCAACACGCG AACAAAAATA AACAGTAGGG ACCATCCGAT TTAAAAAATA
8611 TAATTAGGGA CCAAAAACAT AAATTCCTCC AAACCATAGG GACCATTTCAT GTAATTTACT CTTACTTTTC
8681 GTTTTGTTC TATTTGGGTA ACTATTTTIT TTGTACACAT CTAGGTAACG AACTTGTGTA AGTGTCCCA
8751 TTTAGGATGT GACCTACTAC AACCAGTCAT AATAGTCATA TGTGAACACT TCCAACAAC TTTACTTTA
8821 GGTGTGTACA AAAAAACAAT AGTTACCATG ATGTGAACAT ACTGAAAAAT TAATTACCTT AGCAAGTTAT
8891 TTTCCCATTT AGGTTGTATG GAAACAGTTC CGTGAGACCG TGACTTGGAT GGTAGATAAA TTTAGTAAAC
8961 TTAACCTTTC AATTAACTTA CCTTTTCTT ATTAACCTCA TTTCAACCTA AATTCGTATT CTTGTTTGAA
9031 AGTAAGTTGC ATCTTTATTT TTGTATTATC TTGTGCATA GGATCCTTAG CATCTTTTAA TAATTTATTT
9101 GAAGGTGAAA GATCCAACTA TTTTAAATCT GTTGGCATTT TCCATCATTT GCAACTGTTT CTTGAAAAAA
9171 AAATACCTAA AATCAAAATA ACCATTTTCA AATCCAAAT TATAAGAGAG AATTGTAAAT GGACATGGAA
9241 TCATAAATCA TTAACACAGT TCAGTAAACA AGTTGCTAAT TACATTTCTT GCTGTGCAGA TTGAAATTCT
9311 ATCAGAGAAA GAGACATTAC AAGAAGCCAC TGACAGTATT TCTAATGTTG TATTCCTATC CTGTCTCATG
9381 CACTCTTTTC ATAACCTCCA GAAACTTATA TTGAACAGAG TTAAAGGAGT GGAGGTGGTG TTTGAGATAG
9451 AGAGTGAGAG TCCAACAAGT AGAGAATTGG TAACAACCTA CCATAACCAA CAACAACCTA TTATACCTCC
9521 CAACCTCCAG GAATTGATTC TATGGAATAT GGACAACATG AGTCATGTTG GGAAGTGCAG CAACTGGAAT
9591 AAATCTTCA CTCTTCCAAA ACAACAATCA GAATCCCCAT TCCACAACCT CACAACCATA AAAATTATGT
9661 ATTGCAAAAG CATTAAGTAC TTGTTTTCGC CTCTCATGGC AGAAGTCTT TCCAACCTAA AGCATATCAA
9731 GATAAGAGAG TGTGATGGTA TTGGAGAAGT TGTTCCTAAC AGAGATGATG AGGATGAAGA AATGACTACA
9801 TTTACACTTA CCCACACAAC CACCCTTTG TTCCCTAGTC TTGATCTCT CACTCTAAGT TTCCCTGGAGA
9871 ATCTGAAGTG TATTGGTGA GGTGGTGCCA AGGATGAAGG GAGCAATGAA ATATCTTTCA ATAATACCAC
9941 TGCAACTACT GCTGTTCTTG ATCAATTTGA GGTATGCTTT GTACATATTC AATTATTTAT TTAATTTCTT
10011 TTTTATTTG CAATATTCTA TAAATAATAC ATTTTATACC CACTATACTA AGATAATAAT TACCTAGAGG
10081 GATGGAATGCT ATGACACAGC TGCTACACTT CAGAACTCT AGTAAGGGCA GTTATGGAAG TTCAATAAAA
10151 TGATAATGGC ATCTTTTGAT GGGTAATATA GGCAATTTAA GTTTTATTTT TGTAAAGCA GTATTTAGCA
10221 AGTACTGGCC AGTAGGAGAG GAGAAATACA CCTTTTGTGA AAATCTGGTC ATTGTACCCA GAATTTAGTT
10291 AAATGTAAAC TTTTAGATAT CAGGGGTCAT CAGGTGACAG ATATTGTAGA ATAGAACAAT ATATAATATC
10361 ACCCAAACT ATTTTTCCTA AGGTTATCT GTTAAATATG TGCTTTCTTG TTTTCAINGA ATTTGCAATC
10431 GTATATTTTA GGTGTTAAAG TGATTTTNTC TTCAATAAAT CCGAAAAATTA ATTAATAAAA AAAAAACAAA
10501 AGTACATTTT TGATGTGGAG AGCACTGGTA TCACTTAGTA TATAAAAAGC TTGATTTTGA ATTAACCTTC
10571 TTATACAAAA GTTGTGTATA TAGTTTAAAT AGTTTACAT CATTTTCCA TGTGGTGTG CAGTTGTCTG
10641 AAGCAGGTGG TGTTCCTTGG AGCTTATGCC AATACGCTAG AGAGATGAGA ATAGAATTCT GCAATGCATT
10711 GTCAAGTGTA ATTCATGTT ATGCAGCAGG ACAAATGCAA AAGCTGAAGG AGAGACACG GATTCTCGTA
10781 CGAACGGTTA CGATTCGACT GGCCGTCGTT TTACA

SEQ ID NO: 21

RLGA a.a.

MDVVNAILKPVVETLMVPVKKHIGYLISCRQYMREMGIKMRGLNATRLGVEEHVNRNISNQLEVPAQV
RGWFEEVGKINAKVENFSPDVGSCFNLKVRHGVGKRAASKIIEDIDSVMREHSIIWNDHSIPLGRIDSTK
ASTSIPSTDHHDDEFQSREQTFTALNALDPNHKSHMIALWGMGGVGKTTMMHRLKKVKEKMFNFII
EAVVGEKTDPIAIQSAVADYLGIELNEKTKPARTEKLRKWFVDNSGGKKILVILDDVWQFVDLNDIGLS
PLPNQGVDFKVLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIGVNIVRK
CGGLPIAIKTMACTLRGKSKDAWKNAALLRLEHYDIENIVNGVFKMSYDNLQDEETKSTFLLCGMYPE?FD
ILTEELVRYGWGLKLFKK?YTIGEARTRLNTCIERLIHTNLLMEVDDVRCIKMHDLVRAFVLDMYSKVEH
ASIVNHSNTLEWHADNMHDSCKRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFPPKNFYEEMKLE
VISYDKMKYPLLPSSPQCSVNLRVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDRLPSTIGLKKLR
LLDLTNCYGVRIIDNGVLKLVKLEELYMTVVDGRKAISLTDDNCKEMAERSKDIYALEFFENDAQPK
NMSFEKLQRFQISVGRYLYGDSIKSRHSYENTLKLVEKGELLEARMNELFKKTEVLCLSVGDMDNLEDIE
VKSSSQLLQSSSFNNLRVLVSKCAELKHFFTPGVANTLKKLEHLEVYKCDNMEELJRSRGSEETITTFP
KLKFLSLCGLPKLSGLCDNVKIIELPQLMELELDDIPGFTSIYPMKKFETFSLLEEVLIPKLEKLHVSSM
WNLKEIWPCFNMSEEVKFREIKVSNCDKLVNLFPHKPISLHHLEELKVKNCGSIESLFNIHLDCVGAT
GDEYNNSGVRIIKVISCDKLVNLFPHNPMSILHHLEEEVENCGSIESLFNIDLDCAGAIGQEDNSISLRNI
KVENLGKLR?VWRIKGGDNSRPLVHGFQSVESIRVTKC?KFRNVFTPTTTNFNLGALLEISIDDCGENR
GNDESEESSHEQEIEILSEKETLQEATDSISNVVFPSCLMHSFHNLQKLILNRVKGVEVVFEIESESPTS
RELVTTHHNQQQPILPNLQELILWNMDNMSSHVWKCSNWNKFFTLPKQQSESPFHNLTITIKIMYCKSIKY
LFSPLMAELLSNLKHIKIRECDGIGEVVSNRDDDEEMTTFTSTHTTTTLFPLSLDSTLSFLENLKCIGGG
GAKDEGSNEISFNNTTATTAVLDQFEVCFVHIQLFI.

SEQ ID NO:22

RLG 23

SEQ ID NO: 23

1 AGTTTTTTTTT TTCCCAATA TCCATTTATA TGCGATTTAT TTCTGAAATA ATTTTATCAA AACGCAGGAA
71 ACAATGTAGA ATAATACTGG TATAATTAAT TATATAAAGT TATTAGGCTG AAATCTTGAG GCTACTATAA
141 TTTAATATATC ATAATTGAA AATCATCAAA TTGTATTCCA TGTATATTTA TGTATCAGA TAATTAATAA
211 TATGTGGGCC ACACAAATCC ACATCATCAG ACACCCACC TTATTGTCGG CTACCTCACC ACTTGACATGA
281 TCCCGACATC TTCCCAACCC CACCGACGAC TTGGGGTCTC CTTAATATAT CAATTAATTTT CTGTAAGTAT
351 TTATTTGTGT AAATGTGTAA TGTCAATTTA CCTTTTTTCT AATATATACA GAAACATAAA TTTTAAATGA
421 AATTCAAATCG CGTTTCATTC TTGCATTAAA AAAAAAGACT GTACTGTGT CAATATTTTA CTTATAACCT
491 GATTAATTA TTAAGCGTA ATTGCATAAT TTGCATTAGG TTGTAATTTT GTGTTTTATA GGGAGGGTGA
561 GGGTCACCGG GAATCAAAGC ACTTATGTAA AAGCAGGGGA AATACAAAAA ATTTACTCGA AACAAATTTT
631 ATTCAATTTA AGTGAGATAA TAAATGTTCTG ATTAGATTAT GAGAACTAGG AGATTAAAGT GATATATCCC
701 ATTTAAAGA AATTGCATTA TTAATTTTGG ATCTCTTGAT GATGACAAAA TTAACCTCGT ACAGGTTATA
771 TATCATATAC AAAATGAGTG GCTATGCTTT CGCTTTCCAA AAAGCAATTA TAGTTATACT ACACCTACAA
841 ATTTTAAAG GGGTTAAACA TATCAAAATA CTGTATAAGT AATTATATAA ATATGCATTT AACCTCTAA
911 AGAAAAATGCT ACTAAGCTTG GACCATCTCA GAATTACAAT CATACCCCTC CCCTCAAAAA AGATTGCGAT
981 ATATCAATGTC ATTTGGCATT CATTTCTTTT TCACAATTCA TAGTTCTATT CTCAAAAAAT TCGAGTTCTC
1051 GTATTTGTAA GGAAGATCAG AAGAGACTGT TCACACAGGT ACTCTCTTTT ATTTATTGAT TCACATTCAT
1121 ATATGTTTATT GTTTTCTTGC TTAATGGTTT CGTCAGTCTA ACTGCGCTTG CTGATTTAAA TTTCTTCACT
1191 TTCTTCCACG GATTTTTTAA ATATTAGTTT TGTGAATGAA CAATTTGGTGA AGGAAAGAAA CATGGGAGTC
1261 TTTTCTAAG TAAACCTAGA TACTTAGGTT ATAAGGGTAT ATGCTAAAAT GAACTATGCC CATTCACCTT
1331 TGCCCTTTCT TTTACTTTTT AGTTTTTAGA ATCCAAGTTT TCATATGTAT CTCGATGTGT GAGAAGAATA
1401 GGCATTAGAA AGGTAAAGGA CGTACATAAA ATGTAATAAT TAGTGAATGT TCTTTGATAT CATTAATTTT
1471 ACTCTCATAA AAAGCATATA GATCAAAAC AAATTGCTAC TTGTTAGTGT AACAACTTCG ACTTAATAAT
1541 GTTAATAATC AAGATTCTCT TGATTCAAC TATTTTCTAA CCGAACAAAG TCACATAAAA CTCATATTGC
1611 TTTGAGTCTG AGTGGTTTAT ATTTGGGGTT TTACATTTAA TTTTTTGTGC ATGAATGTGA AAATAGACTG
1681 CTTATTGATT CTTTGTGTTT CATTGAGTTG ATTTTCATTA TTACTACCTT ACAAATTGCT CAGTGATAGA
1751 TTTCCATTA TTTGCTAATT CGGTGCTTTC TAAATATGTA GGAGCTACTA AAAGCAAAAA TATCGAGCAA
1821 TGTCGGACCC AACGGGGATT GCTGGTGCCA TTATTAACCC AATTGCTCAG ACGGCTTGG TTCCCGTTAC
1891 GGCCCATGTA GGCTACATGA TTCTCTGCG AATAATATGT AGGGTCATGC AGATGAAAAT GACAGAGTTG
1961 AATACCTCAA GAATCAGTGT AGAGGAACAC ATTAGCCGGA ACACAAGAAA TCATCTTCAG TTCCATCTCA
2031 AACTAAGGAA TGGTTGGACC AAGTAGAAGG GATCAGAGCA AATGTGGAAA ACTTTCCGAT TGATGTCATC
2101 ACTTGTGTGA GTCTCAGGAT CAGGCACAAG CTGTGACAGA AAGCNTTCAA GATAACTGAG CAGATTGAAA
2171 GTCTAATCGA ACAACTCTCC CTGATCAGTT GGACTGATGA TCCAGTTCYT CTAGGAAGAG TTGGTTCCAT
2241 GAATGCATCC ACCTCTGCAT CATTAAGTGA TGATTTCCCA TCAAGAGAGA AAACCTTTTAC ACAAGCACTA
2311 ATAGCTCTCG AACCCAACCA AAAATTCAC ATGGTAGCCT TGTGTGGGAT GGGTGGAGTG GGAAGACTA
2381 GAATGATGCA AAGGCTGAAG AAGGCTGATG AAGAAAAGAA ATTTGTTAAT TATATTGTTG GGGCAGTTAT
2451 AKGGGAAAAG ACGGACCCCT TTGCCATTCA AGAAGCTATA GCAGATTACC TCGGTATACA ACTCAATGAA
2521 AAAACTAAGC CAGCAAGAGC TGATAAGCTT CGTGAATGGT TCAAAAAGAA TTCAGATGGA GGTAAAGACTA
2591 AGTTCCCTCAT AGTACTTGAC GATGTTTGGC AATTAGTTGA TCTTGAAGAT ATTGGGTTAA GTCCTTTTC
2661 AAATCAAGGT GTCGACTTCA AGGTCTTGTT GACATCACGA GACTCACAAG TTTCGACTAT GATGGGGTT
2731 GAAGCTAATT CAATTATTAA CGTGGGCCTT CTAACCTGAAG CAGAAGCTCA AAGTCTGTTT CAACAATTTG
2801 TAGAACTTTC TGAGCCCGAG CTCCAGAAGA TAGGAGAGGA TATCGTAAGG AAGTGTTCGG GTCTACCTAT
2871 TGCCATAAAA ACCATGGCAT GTWCTCTTAG AAATAAAGA AAGGATGCAT GGAAGGATGC ACTTTCCGCG
2941 ATAGAGCACT ATGACATTCA CAATGTTGCG CCCAAAGTCT TTGAAACGAG CTACCACAAT CTCCAAGAAG
3011 AGGAGACTAA ATCCACTTTT TTAATGTGTG GTTTGTTTCC CGAAGACTTC GATATTCTTA CTGAGGAGTT
3081 GATGAGGTAT GGATGGGGCT TGAAGCTATT TGATAGAGTT TATACGATTA GAGAAGCAAG AACAGGCTC
3151 AACACCTGCA TTGAGCGACT GGTGCAGACA AATTGTGTTA TTGAAAGTGA TGATGTTGGG TGTGTCAAGA
3221 TGCATGATCT GGTCCGTGCT TTGTGTTTGG GTATGTTTTC TGAAGTCGAG CATGCTTCTA TTGTCAACCA
3291 TGGTAATATG CCTGGGTGGC CTGATGAAAA TGATATGATC GTGCATCTT GCAAAAAGAT TTCATTAAAC
3361 TGCAAGGGTA TGATTGAGAT TCCAGTAGAC CTCAAGTTTC CTAAACTAAC GATTTTGAAG CTTATGCATG
3431 GAGATAGTC GCTAAGGTTT CCTCAAGACT TTTATGAAGG AATGGAAAAG CTCCATGTTA TATCATACCA
3501 TAAAAAGAAG TACCAATGTC TTCTTTTGGC ACCTCGATGC TCCACCAACA TTGGGGTGCT TCATCTCACT
3571 GAATGTCAT TAAAGATGTT TGATTGCTCT TCTATCGGAA ATCTATCGAA TCTGGAAGTG CTGAGCTTTG
3641 CAAATCTCA CATGAAATGG TTACCTTCCA CAGTCAGAAA TTTAAAGAAG CTAAGGTTAC TTGATCTGAG
3711 ATTTTGTGAT GGTCTCCGTA TAGAACAGGG TGTCTTGAAA AGTTTTGTCA AACTTGAAGA ATTTTATATT
3781 GGAGATGAT CTGGGTTTAT AGATGATAAC TGCAATGAGA TGGCAGAGCG TTCTTACAAC CTTTCTGCTA
3851 TAGAATTCGG TCTCTTAAAT AACAAAGCTG AAGTGAAAAA TATGTCAATT GAGAATCTTG AACGATTCAA
3921 GATCTCAGTG GGATGCTCTT TTGATGAAAA TATCAATATG AGTAGCCACT CATACGAAAA CATGTTGCAA
3991 TTGGTGACCA ACAAAGGTGA TGTATTAGAC TCTAAACTTA ATGGGTTATT TTGAAAAACA GAGGTGCTTT
4061 TTTTAAATGT GCATGGCATG AATGATCTTG AAGATGTTGA GGTGAAGTCG ACACATCTTA CTCAGTCTC

RLG2B cont.

SEQ ID NO:23

4131 TTCATTCTGC AATTTAAAG TTCATTAT TTCAAAGTGT GTAGAGTTGA GATACCTTTT CAAACTCAAT
4201 CTTGCAACA CTTTGTCAAG ACTTGAGCAT CTAGAAGTTT GTGAATGTGA GAATATGGAA GAACTCATAC
4271 ATACTGGAAT TGGGGGTGT GGAGAAGAGA CAATTACTTT CCCTAAGCTG AAGTTTTTAT CTTTGAGTCA
4341 ACTACCGAAG TTATCAAGTT TGTGCCATAA TGTCACATA ATTGGGCTAC CACATCTCGT AGACTTGATA
4411 CTTAAGGGCA TTCCAGGTTT CACAGTCATT TATCCGCAGA ACAAGTTGCG AACATCTAGT TTGTTGAAGG
4481 AAGGGGTAGA TATATGTTCT TTATGTTAAT ACAATTTAAA TAATATTTTC AACCATAATT TCATAATATA
4551 TCTGTAATTT GATTGTATGA TGTGTTATTG TTTATATGTG GCTATTAAGG GATGATTATT TTGCAGGTG
4621 TGATTCCTAA GTTGGAGACA CTTCAAATTG ATGACATGGA GAACTTAGAA GAAATATGGC CTTGTGAAT
4691 TAGTGGAGGT GAGAAAGTTA AGTTGAGAGC GATTAAAGTG AGTAGCTGTG ATAAGCTTGT GAATCTATT
4761 CCGCGCAATC CCATGTCCTT GTTGCAATCAT CTGAAGAGC TTACAGTCCA GAATGCGGT TCCATTGAGT
4831 CGTTATTCAA CATTGACTTG GATTGTGTG GTGCAATTGG AGAAGAAGAC AACAGAGGCC TCTTAAGAAG
4901 CATCAACGTG GAGAATTTAG GGAAGCTAAG AGAGGTGTGG AGGATAAAG GTGCAGATAA CTCTGATCTC
4971 ATCAACGGTT TTCAAGCTGT TGAAGCATAA AAGATTGAAA AATGTAAGAG GTTTAGAAAT ATATTACAC
5041 CTATCACCOC CAATTTTTAT CTGGAGGCAC TTTTGGAGAT TCAGATAGAA GGTGCGGAG GAAATCAGCA
5111 ATCAGAAGAG CAGGTAACGC TTCAATTTT ACITTCCTAA TTAATTAAGG ACTAAGCTCC TGTTTTTTGA
5181 ATAATAAGA GGTGGATGA CTAACTTGG GCATCACAAT TGCAACAAA TGTTACAAAC CATGAAACGT
5251 TCAAACCTAT TCTTGAATTA AGGTTTCAAT ACAAGTCATT TAAAAATATG GCTTAAATTT TTTTATATT
5321 TATGTATCAA CATGATTTTT CATTAGAGAT CATTATTATA ATAGTAAGTT TAAAGCAATT TAAATCAGAA
5391 CTAATTTCTA CTTTAGCTAA TAAATCGTTA TAAATGTAAA TAATTACTTT TTAGTGAAAT AAGCAACGGA
5461 TTTAATAAGT TAACAACCTA AATGTCAATT CCTAACAAA AAAACTTTTG TTCAGAAAAA CCGCAATTC
5531 AGATAACTAA AATAAAAAA TTTGACATTC ACTAAGAGCA TTTTTTTTTC TAAATATGAT TGCAAAATGAA
5601 TAAAACTTAA ATTTATACAG AAAATCTTT TATATATGTT ATACAAAAT TACAAAATGA AATTGGATAT
5671 GTTAATTAA GGTTTATAAT TCTGGTATCA CAAAGGGATA TATAATAAAA TATTATTTTC TGTAGTCATT
5741 TGTAAATGTA CTAGTTTATA ACCCGTGGGA ACCATGAGTT CTAAATTTAG TTAACCTTTC ATAATAAAA
5811 TTTTATAATTA TTATTATTAT TAAATAAATT ATTAATTAA AGATATATCA AAAAATTTAA GTTATTATA
5881 CTTCAAAATT AACATATAAT TAGAAAAAT ATGATCATAA CTTCTGCACT CTCTTTGTAT AAATGCAGAG
5951 AAGCTATTAG TATATTCTA ATCAAGTCCA AACCTAATGA AGCCTATATA ATTTTGTGAA AACTCAATTA
6021 GCATTAGGTT TTAAGAGTCA CCAAAATCCA AGAATAATCC AATGCTTTCA TTACCCTAT TCCAAAATA
6091 TTTCTTAGT TTAATGAAA TGAACAACA CATTCAACT AATTGTTGCT TATTAAACCA AAGACCCATT
6161 ACTTAGCCAA GAGTTTAA CAAAAAAAT ACATTCACT ATCATTATT ATGACTAGAT ATATATGAAC
6231 GTTCTTCTAA GTTCTTCAAG TCTCCAGTGT CTCTTAATAA TGCAAGGAG AGCCATTAAA TTCGTATGTA
6301 CCAAGTTATT CAAGAAATTA CATCCAAGTC AACCAAGAG AAGTTTAGCC TAGCATGGCT AAACCTAAGA
6371 AACTAAAAA AGGATTAGAA GTACCAACA TGTAGTAAGA ATCAGAGTAA AAGATGATGT TGTCTTGAT
6441 GTTCTTCTAA GTTCTTCAAG TCTCCAGTGT CTCTTAATAA TGCAAGGAG AGCCATTAAA TTCGTATGTA
6511 TTGATCCCTT CAAAAGCTGC ACCAACCTCC CTTAAATAAC ACTCAAAGCA AAAATGACAA AATGCCCTGA
6581 AGGACCCCTT GTGGGTGCTT TCGCGGGGTG GAGCTGCATA CGAAAGGTCT TTGGTCTTTG TGAGGGTGAT
6651 GTTGTGCGG ATAGCTTGTG GCATGCTTCC GCGCGGTTC CACACATGTG CACAGGTGAT GCATGGTGTG
6721 TCGGTCTTTC AGTTTGTAGC CTCCGATGCT TAGTCCACTT GCGCCCAATC GAGTCCAATC AGCTTATAAC
6791 CCATTTTCT TCAAGTTATC TTCAAGTTAA GCCCAATTTG GCTTCTCCAA ATCATCCATA ACTTCACAGA
6861 ATGCGCGGTT CATCTTAATC CCGGATGCAC AATTATTCTC CGGTCTTCAT TTAAAGCAAG ATACCACCTT
6931 CTTCTAGCTT CATCCATCAA TAGTACACTT CATGTATCAT CTCTACTAGT TATTTAGTCC ACAAATCCCT
7001 GTTGTCTTCC AAATTTAATT ATCTCAATTA GTTCCCGGT COGCTACTTT CCTTAAATTT TGGAAATTAAG
7071 CTCAGAGAAA TATTAAGTAC CCGAAATGGT CATAAAATTA ACAAAGGA AAATGCATGA AGATTAACCTA
7141 AATGATGAAC GAAATATGCT AAAATAGACT ATAAATAGAA GTAAATATAA TGAAATATAT GCCTCCGAC
7211 CACCTTATG GCTTGTAGTC CACCCACCTT TCATTCCTTG TACCAATATG GGATGGAAC ATCATTAAAT
7281 AAGCCAAAA GCTAACATAT AAGGGTTTAG TGACAAAGGT AAGTACTAAA GATGAAAATA ATCCATTTT
7351 CTTGTTTTTA CACAACACAC ACATAGGGGC AGACGTAGGA TTTCAAAGTA CAGATTGTTG GTGGCACATA
7421 AGTGTGCTG GTGACATTTT TTTTCTCTT TTACGTGGTG GCACAACAGT AGGAAAAACG AAAAATTCGA
7491 AATTTTTTT AATTTGTCTT AAAAAAACA GGGGTGTGTG GTGCCACTAT GGCAACAAA GTTGAACCTG
7561 CTTACCGCGG CACACACACA CACACACATA GAGAGAGAGA GAGAGAGAGA GAGAGAGAGA AAGAAAGAAA
7631 GAGAGAGAGA GTTTGGGATG TGATACTTCT TTTAGGAAAA TGGAGTTATA TCTTTGATAT TGTATTTTTT
7701 TAAATGAATT TATNTATTTA ATCATTTTAG TTTATAAGTT NTATTTATIN GGNATGAAA AAAAAGTCT
7771 TTTATCAATT GGATTTAACA TAAAAATCCA ACAATATTAA TCAAAAGAC CAAACATGTG GACAAATTAT
7841 TATATAATTA ATTCACAATA GTCTTTAGGA ATAGTATTAT ATATATAATT AATTCTCAAT GGTCTTAGGA
7911 ATAGTAAGTT CTTATATTTT AAACCTTTTC CACAATTTCT TGCTTACTTT GACACTTTTC CTTCTTAAC
7981 TTACATATAT ATATATATTA AAGCGCAAAG GTCATAGGAA TATAATATTT TCTATTATTC TACGTTTTGC
8051 CACAAAAGTT TGAACACTTT GCCACTTTTT GTCCCTCTTT AACCTTTTCA ATGTTTTGCG ACAAAAGTTC
8121 CAAAACCTTG CCACTTTGAT CATTCCTCAA CTTTTCACCG CATTAGTTTG TGGAGTTGGC AGTTTTGCTC
8191 CCTCTAAGTT CGATATTCTC TACTGCTAGC CAAAAGGGT TCCAGAGTTT CACACTTTTG GTCCCTGACA

RLG 2B cont.

8261 GTAACCAAT GTGAGATGTC AAATTTTTC CACATTAGTT TGTGGAGTTG TCCCTTTTGG TCCCCCACA
8331 TTGCATATTC TACTATACGA TCTTATTTTT CTCAAATAAC AACACGTATA TTTTATC:CT AATTGGAAAA
8401 AGAGTTTTTAA AA:AAATAAC GACTAGG::: G:GC:GAGTT TTTTTT:ACA AGTTTGTATC AAATCATATC
8471 AAAATTTAAG GTGGAACGGT GACCACATTA ACCAGAAATG TAATTTATTC TTTGATTTTG ATAATTTTTTA
8541 ATATTTTGTG GTGATCTATG TATTTAAAAG TAAACAACAA AGAACATAAT CCAAAACCCCT AAATTTGCAAG
8611 TCTCGCCCAA TTTCTCTATC ACTAGTCCCTC ACTTACGATG GCGTTACGTC GCTCTCTCAC TGCTTACAAC
8681 CCTTTGTGTC TACTCATTAC AATAACGAAA AGTTGAATAT CCATATATTT ATTTGGATGT GGAATTGAAC
8751 GAATCTCGTC AAAATTTTGA TTTTGTGAT GGATTTGAGT AGAAGTTTGG GCAGAACGGG AATGATGGTC
8821 TGCAAGTGGT TATAAACTTG ATTCTGAGTT ATTACTATAT ATGTAGCCTC TTTACAACGA CCAAGGTTTC
8891 TTCCAGGTAC CATTTGATCT TTTTAGAACT TAGTTTTCTG AAACACCCCTG ATTTGGATCA AATATCACCA
8961 ACAACTCTTA AAAACTTGAT TAATCAATTG TTTTCTTCAT CTGTATAACA AGTGGAAATGA TTTTCTACTT
9031 AGATTAACCTT GAAAAAAAAG GTCCATGTGC GTCTGGTGGG TCTGGTAAAT GAAGATGGAA GGGAGAGCTG
9101 ACTTTAAAGA CACAAACACG TCACCATATC TCTTATTTTA TTTTAAATTT GCTTTTGTG TATTTTCTTT
9171 TTTCTCTTTT CTTTCTTTCT TGATCTCCAG ATGGTATGTG GTGTGGATAA TTTACACCTA GAGATTGGGA
9241 ACGATGGGAA GGGGTCTGTG ATTTATGGCT GGCCGAGTTT TACTTATTA CTCAATTTCA ACCTAAATTC
9311 TGATTCCTGT TTGAAAATAA GTTGCACTCT TATTTTGTG TTATCTGTGT GCATAGGATC CTTAGCATCT
9381 TTTAATAATT TATTTGAAGG TGAAAGATCC AACTATTTTT TAGCTGTGG CATTTTCCAT CATTGCAAC
9451 TGTTCCTTGA AAAAAAATA CCTAAATAA AAATAACCAT TTTCAAATCC AAAATTATAA GAGAGAATTG
9521 TAAATGGACA TGGAAATCATA AATCATTAAC ACAGTTTCAAGT AAACAAGTTG CTAATTACAT TCTTGCTGT
9591 GCAGATTGAA ATTCTATCAG AGAAAGAGAC ATTACAAGAA GCCACTGGCA GTATTTCAAA TCTTGTATTC
9661 CCATCCCTGC TCATGCATCT TTTTCATAAC CTCCTGTGTC TTACATTTGA TAATATGAA GAGATGGAGG
9731 TGGTATTTGA GATAGAGAT GAGAGTCCAA CATGTAGAGA ATGGTAACA ACTCGCAATA ACCAACACA
9801 GCCTATTTATA CTTCCCTACC TCCAGGATTT GTATCTAAGG AATATGGACA ACACGAGTCA TGTGTGGAAG
9871 TGCAGCACT GGAATAAATT CTTCACTCTT CCAAAACAAC AATCAGAATC CCCATTCCAC AACCTCACA
9941 CCATAAATAT TCTTAAATGC AAAAGCATT AATACCTTGT TTCGCTCTC ATGGCAGAAC TTTCTTCCAA
10011 CCTTAAAGGAT ATCCGGATAA GTGAGTGTGA TGGTATTAAA GAAGTTGTTT CAAACAGAGA TGATGAGGAT
10081 GAAGAAATGA CTACATTTAC ATCTACCCAC ACAACCACCA CTTTGTTCCT TAGTCTTGAT TCTCTCACTC
10151 TAAGTTTCTT GGAGAATCTG AAGTGTATTG GTGGAAGTGG TGCCAAGGAT GAGGGGAGCA ATGAAATATC
10221 TTTCAAATAA ACCACTGCAA CTACTGCTGT TCTTGATCAA TTTGAAGTAT GCTTTGTACA TATTCATTA
10291 TTTATTTAAT TTCTTTTAT ATTTGCAATA TTCTATAAAT AATACATTTT ATACCCACTA TACTAAGATA
10361 ATAATTACCT AGAGGGATGG ATGCTATGAC ACAGCTGCTA CACTTCAGAA ACTCTARTAA GGGCAGTTAT
10431 GGAAGTTCAA TAAAATGATA ATGGCATCTT TTGATGGGTA ATATAGGCAA TTTAAGTTT ATTTCTGTTA
10501 AAGCAGTATT TAGCAAGTAC TGGCCAGTAG GAGAGGAGAA TATCACCTTT TGTGAAAATC TGGTCATTGT
10571 ACCCAGAATT TAGTTAAATG TAACATTTTA GATATTAGGG GTTATCAGGT GACAGATATT GTAGAATAGA
10641 ACAATATGTA ATATTACCCA AAACATTTTT TTCTAAGGTT GCTCTGTAA ATATGTGCTT TCTTGATTTT
10711 ATTGAATTTG CATTCCTATA TTTTAGGTGG TAAAGTGATT GTCTCTTCAA TAAATCCCGA AATTTTTTAA
10781 TTAATAAAAA AAAAAACAA AGTAAATTTT TGATATGGAG AGCACTGGTA TCATTTAGTA TATAAAAAAC
10851 AGATTTTGA TTAAGTTTCT TATATAAAAG CTGTGTATAT AGTTTAAATTA GTTTTACATC ATTTTCCAT
10921 GTGGTGTGTC AGTTGTCTGA AGCAGGTGGT GTTCTTGTGA GCTTATGCCA ATACGCTAGA GAGATAAAAA
10991 TAGGCAACTG CCATGCATTG TCAAGTGTGA TTCCATGTTA TGCAGCAGTA CAAATGCAGA AAGCTT

SEQ ID NO: 23

RLG2B a.u.

MSDPTGIAGAIINPIAQTALVPVTDHVGYMISCRKYVRVMQMKMTELNTSRISVEEHISRNTRNHLQIP
SQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLGGKAFKITEQIESLTRQLSLISWTDDPV?LGRVG
SMNASTSASLSDDFPSREKTFTQALIALEPNQKFHMVALCGMGGVGKTRMMQRLKKA?EEKLFFNYIV
GAVI?EKTDPFAIQEAIADYLGQQLNEKTKPARADKLREWFKKNSDGGGKTKFLIVLDDVWQLVDLEDIGL
SPFPNQGVDFKVLTSRDSQVCTMMGVEANSIINVGLL TEAEASLFFQQFVETSEPELQKIGEDIVRKC
CGLPIAIKTMAC?LRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQEEETKSTFLMCGLFPEDFDI
PTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGCVKMHDLVRAFLVGMFSEVEH
ASIVNHGNMPGWPDENDMIVHSCKRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLRFPQDFYEGMEKL
HVISYDKMKYPLLPLAPRCSTNIRVLHLTECSLKMFCSSIGNLSNLEVLSFANSHIEWLPSTVRNLKKL
RLDLRFCDGLRIEQGVLSFVKLEEFYIGDASGFIDNDCNEMAERSYNLSALEFAFFNNKAEVKNMSFE
NLERFKISVGCSDENINMSSHSYENMLQLVTNKGDVLD SKLNGFLKTEVLFLSVHGMNDLEDVEVKS
THPTQSSSFCNLKVLISKVELRYLFKLNLANLTSRLEHLEVCECENMEELIHTGIGGCGEETITFPKLF
LSLSQLPKLSSLCHNVNIIIGLPHLVDLILKGIPGFTVIYPQNKLR TSSLLKEGVVIPKLETLOIDDMENLEE
IWPCELSGGGEKVKLRAIKVSSCDKLVNLFPRNPMSLLHHLEELTVENCGSIESLFNIDLDCVGAIGEEDN
KSLLRSINVENLGKLREVWRIKGADNSDLINGFQAVESIKIEKCKRFRNIFTPTANFYLEALLEIQIEGCG
GNHESEEQVTLISLS

SEQ ID NO: 24

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

[illegible]

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

RLG2A	810	820	830	840	850	860	870	880	890	900
RLG2B	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2C	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2D	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2E	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2F	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2G	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2H	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2I	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2J	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2K	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2L	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2M	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2N	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2O	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2P	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2Q	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2R	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2S	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2T	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2U	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2V	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2W	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2X	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2Y	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2Z	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2A	910	920	930	940	950	960	970	980	990	1000
RLG2B	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2C	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2D	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2E	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2F	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2G	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2H	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2I	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2J	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2K	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2L	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2M	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2N	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2O	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2P	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2Q	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2R	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2S	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2T	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2U	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2V	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2W	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2X	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2Y	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC
RLG2Z	GC	TC	AA	CA	CG	TC	GA	TC	GA	TC

[illegible]

	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500
RLG2A	CTTAAAAAATTTGGTCAAACTGGAGGCTCTATAT	GCACAGTG	---	GTT	---	CAT	---	CCAGGTGCMAAGG	---	GCA
RLG2B	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GACATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1475
RLG2C	CTTCAAAAATTTGGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1477
RLG2D	CTTCAAAAATTTGGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1479
RLG2E	CTTCAAAAATTTGGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1481
HLG2F	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1483
RLG2G	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1485
RLG2H	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1487
RLG2I	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1489
RLG2J	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1491
RLG2K	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1493
RLG2L	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1495
RLG2M	CTTCAAAAATTTGTCAAACTGACGATTTTATATTTG	GAATGATCATCTGGT	TTATATGAT	---	TATATGAT	---	TATATGAT	---	TTACACCTT	1497



	1510	1520		
RIG2A	-----TTA-	-SEQ	ID NO: 27	1439
RIG2B	TTCGATTAGAAATCGCGTTTCTTA	-SEQ	ID NO: 28	1500
RIG2C	-----	-SEQ	ID NO: 29	1417
RIG2D	-----TG	-SEQ	ID NO: 30	1412
RIG2E	-----	-SEQ	ID NO: 31	1432
RIG2F	-----TA-	-SEQ	ID NO: 32	1438
RIG2G	-SEQ ID NO: 33			1466
RIG2H	-SEQ ID NO: 34			1420
RIG2I	-SEQ ID NO: 35			1442
RIG2J	TCGAATGGCGAGACCTTTTTCACA	-SEQ	ID NO: 36	1474
RIG2K	TCGATTCACCAAGGC	-SEQ	ID NO: 37	1437
RIG2L	-----	-SEQ	ID NO: 38	1398
RIG2M	-----A-	-SEQ	ID NO: 39	1441

1310
1320
-----TTA - SEQ ID NO: 27
TCTGCATTAGAAATCGCGTTCTTTA - SEQ ID NO: 28
-----SEQ ID NO: 29
-----TG - SEQ ID NO: 30
-----SEQ ID NO: 31
-----TA - SEQ ID NO: 32
ED NO: 33
ED NO: 34
ED NO: 35
TGGAAATGCGAGACCTTTTTCACA - SEQ ID NO: 36
TCGATATCAAAAGC - SEQ ID NO: 37
-----SEQ ID NO: 38
-----A - SEQ ID NO: 39

SEQ ID NO:

40
 10 20 30 40 50 60 70 80 90 100
 110 120 130 140 150 160 170 180 190 200
 210 220 230 240 250 260 270 280 290 300
 310 320 330 340 350 360 370 380 390 400
 410 420 430 440 450 460 470 480 490 500
 510 520 530 540 550 560 570 580 590 600
 610 620 630 640 650 660 670 680 690 700
 710 720 730 740 750 760 770 780 790 800
 810 820 830 840 850 860 870 880 890 900
 910 920 930 940 950 960 970 980 990 1000
 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900
 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100
 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200
 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300
 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500
 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600
 2610 2620 2630 2640 2650 2660 2670 2680 2690 2700
 2710 2720 2730 2740 2750 2760 2770 2780 2790 2800
 2810 2820 2830 2840 2850 2860 2870 2880 2890 2900
 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100
 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200
 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300
 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400
 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500
 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800
 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900
 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000
 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100
 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300
 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400
 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500
 4510 4520 4530 4540 4550 4560 4570 4580 4590 4600
 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700
 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900
 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000
 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100
 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200
 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300
 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
 5410 5420 5430 5440 5450 5460 5470 5480 5490 5500
 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600
 5610 5620 5630 5640 5650 5660 5670 5680 5690 5700
 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800
 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900
 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000
 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100
 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200
 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300
 6310 6320 6330 6340 6350 6360 6370 6380 6390 6400
 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500
 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600
 6610 6620 6630 6640 6650 6660 6670 6680 6690 6700
 6710 6720 6730 6740 6750 6760 6770 6780 6790 6800
 6810 6820 6830 6840 6850 6860 6870 6880 6890 6900
 6910 6920 6930 6940 6950 6960 6970 6980 6990 7000
 7010 7020 7030 7040 7050 7060 7070 7080 7090 7100
 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200
 7210 7220 7230 7240 7250 7260 7270 7280 7290 7300
 7310 7320 7330 7340 7350 7360 7370 7380 7390 7400
 7410 7420 7430 7440 7450 7460 7470 7480 7490 7500
 7510 7520 7530 7540 7550 7560 7570 7580 7590 7600
 7610 7620 7630 7640 7650 7660 7670 7680 7690 7700
 7710 7720 7730 7740 7750 7760 7770 7780 7790 7800
 7810 7820 7830 7840 7850 7860 7870 7880 7890 7900
 7910 7920 7930 7940 7950 7960 7970 7980 7990 8000
 8010 8020 8030 8040 8050 8060 8070 8080 8090 8100
 8110 8120 8130 8140 8150 8160 8170 8180 8190 8200
 8210 8220 8230 8240 8250 8260 8270 8280 8290 8300
 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400
 8410 8420 8430 8440 8450 8460 8470 8480 8490 8500
 8510 8520 8530 8540 8550 8560 8570 8580 8590 8600
 8610 8620 8630 8640 8650 8660 8670 8680 8690 8700
 8710 8720 8730 8740 8750 8760 8770 8780 8790 8800
 8810 8820 8830 8840 8850 8860 8870 8880 8890 8900
 8910 8920 8930 8940 8950 8960 8970 8980 8990 9000
 9010 9020 9030 9040 9050 9060 9070 9080 9090 9100
 9110 9120 9130 9140 9150 9160 9170 9180 9190 9200
 9210 9220 9230 9240 9250 9260 9270 9280 9290 9300
 9310 9320 9330 9340 9350 9360 9370 9380 9390 9400
 9410 9420 9430 9440 9450 9460 9470 9480 9490 9500
 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600
 9610 9620 9630 9640 9650 9660 9670 9680 9690 9700
 9710 9720 9730 9740 9750 9760 9770 9780 9790 9800
 9810 9820 9830 9840 9850 9860 9870 9880 9890 9900
 9910 9920 9930 9940 9950 9960 9970 9980 9990 10000
 10010 10020 10030 10040 10050 10060 10070 10080 10090 10100
 10110 10120 10130 10140 10150 10160 10170 10180 10190 10200
 10210 10220 10230 10240 10250 10260 10270 10280 10290 10300
 10310 10320 10330 10340 10350 10360 10370 10380 10390 10400
 10410 10420 10430 10440 10450 10460 10470 10480 10490 10500
 10510 10520 10530 10540 10550 10560 10570 10580 10590 10600
 10610 10620 10630 10640 10650 10660 10670 10680 10690 10700
 10710 10720 10730 10740 10750 10760 10770 10780 10790 10800
 10810 10820 10830 10840 10850 10860 10870 10880 10890 10900
 10910 10920 10930 10940 10950 10960 10970 10980 10990 11000
 11010 11020 11030 11040 11050 11060 11070 11080 11090 11100
 11110 11120 11130 11140 11150 11160 11170 11180 11190 11200
 11210 11220 11230 11240 11250 11260 11270 11280 11290 11300
 11310 11320 11330 11340 11350 11360 11370 11380 11390 11400
 11410 11420 11430 11440 11450 11460 11470 11480 11490 11500
 11510 11520 11530 11540 11550 11560 11570 11580 11590 11600
 11610 11620 11630 11640 11650 11660 11670 11680 11690 11700
 11710 11720 11730 11740 11750 11760 11770 11780 11790 11800
 11810 11820 11830 11840 11850 11860 11870 11880 11890 11900
 11910 11920 11930 11940 11950 11960 11970 11980 11990 12000
 12010 12020 12030 12040 12050 12060 12070 12080 12090 12100
 12110 12120 12130 12140 12150 12160 12170 12180 12190 12200
 12210 12220 12230 12240 12250 12260 12270 12280 12290 12300
 12310 12320 12330 12340 12350 12360 12370 12380 12390 12400
 12410 12420 12430 12440 12450 12460 12470 12480 12490 12500
 12510 12520 12530 12540 12550 12560 12570 12580 12590 12600
 12610 12620 12630 12640 12650 12660 12670 12680 12690 12700
 12710 12720 12730 12740 12750 12760 12770 12780 12790 12800
 12810 12820 12830 12840 12850 12860 12870 12880 12890 12900
 12910 12920 12930 12940 12950 12960 12970 12980 12990 13000
 13010 13020 13030 13040 13050 13060 13070 13080 13090 13100
 13110 13120 13130 13140 13150 13160 13170 13180 13190 13200
 13210 13220 13230 13240 13250 13260 13270 13280 13290 13300
 13310 13320 13330 13340 13350 13360 13370 13380 13390 13400
 13410 13420 13430 13440 13450 13460 13470 13480 13490 13500
 13510 13520 13530 13540 13550 13560 13570 13580 13590 13600
 13610 13620 13630 13640 13650 13660 13670 13680 13690 13700
 13710 13720 13730 13740 13750 13760 13770 13780 13790 13800
 13810 13820 13830 13840 13850 13860 13870 13880 13890 13900
 13910 13920 13930 13940 13950 13960 13970 13980 13990 14000
 14010 14020 14030 14040 14050 14060 14070 14080 14090 14100
 14110 14120 14130 14140 14150 14160 14170 14180 14190 14200
 14210 14220 14230 14240 14250 14260 14270 14280 14290 14300
 14310 14320 14330 14340 14350 14360 14370 14380 14390 14400
 14410 14420 14430 14440 14450 14460 14470 14480 14490 14500
 14510 14520 14530 14540 14550 14560 14570 14580 14590 14600
 14610 14620 14630 14640 14650 14660 14670 14680 14690 14700
 14710 14720 14730 14740 14750 14760 14770 14780 14790 14800
 14810 14820 14830 14840 14850 14860 14870 14880 14890 14900
 14910 14920 14930 14940 14950 14960 14970 14980 14990 15000
 15010 15020 15030 15040 15050 15060 15070 15080 15090 15100
 15110 15120 15130 15140 15150 15160 15170 15180 15190 15200
 15210 15220 15230 15240 15250 15260 15270 15280 15290 15300
 15310 15320 15330 15340 15350 15360 15370 15380 15390 15400
 15410 15420 15430 15440 15450 15460 15470 15480 15490 15500
 15510 15520 15530 15540 15550 15560 15570 15580 15590 15600
 15610 15620 15630 15640 15650 15660 15670 15680 15690 15700
 15710 15720 15730 15740 15750 15760 15770 15780 15790 15800
 15810 15820 15830 15840 15850 15860 15870 15880 15890 15900
 15910 15920 15930 15940 15950 15960 15970 15980 15990 16000
 16010 16020 16030 16040 16050 16060 16070 16080 16090 16100
 16110 16120 16130 16140 16150 16160 16170 16180 16190 16200
 16210 16220 16230 16240 16250 16260 16270 16280 16290 16300
 16310 16320 16330 16340 16350 16360 16370 16380 16390 16400
 16410 16420 16430 16440 16450 16460 16470 16480 16490 16500
 16510 16520 16530 16540 16550 16560 16570 16580 16590 16600
 16610 16620 16630 16640 16650 16660 16670 16680 16690 16700
 16710 16720 16730 16740 16750 16760 16770 16780 16790 16800
 16810 16820 16830 16840 16850 16860 16870 16880 16890 16900
 16910 16920 16930 16940 16950 16960 16970 16980 16990 17000
 17010 17020 17030 17040 17050 17060 17070 17080 17090 17100
 17110 17120 17130 17140 17150 17160 17170 17180 17190 17200
 17210 17220 17230 17240 17250 17260 17270 17280 17290 17300
 17310 17320 17330 17340 17350 17360 17370 17380 17390 17400
 17410 17420 17430 17440 17450 17460 17470 17480 17490 17500
 17510 17520 17530 17540 17550 17560 17570 17580 17590 17600
 17610 17620 17630 17640 17650 17660 17670 17680 17690 17700
 17710 17720 17730 17740 17750 17760 17770 17780 17790 17800
 17810 17820 17830 17840 17850 17860 17870 17880 17890 17900
 17910 17920 17930 17940 17950 17960 17970 17980 17990 18000
 18010 18020 18030 18040 18050 18060 18070 18080 18090 18100
 18110 18120 18130 18140 18150 18160 18170 18180 18190 18200
 18210 18220 18230 18240 18250 18260 18270 18280 18290 18300
 18310 18320 18330 18340 18350 18360 18370 18380 18390 18400
 18410 18420 18430 18440 18450 18460 18470 18480 18490 18500
 18510 18520 18530 18540 18550 18560 18570 18580 18590 18600
 18610 18620 18630 18640 18650 18660 18670 18680 18690 18700
 18710 18720 18730 18740 18750 18760 18770 18780 18790 18800
 18810 18820 18830 18840 18850 18860 18870 18880 18890 18900
 18910 18920 18930 18940 18950 18960 18970 18980 18990 19000
 19010 19020 19030 19040 19050 19060 19070 19080 19090 19100
 19110 19120 19130 19140 19150 19160 19170 19180 19190 19200
 19210 19220 19230 19240 19250 19260 19270 19280 19290 19300
 19310 19320 19330 19340 19350 19360 19370 19380 19390 19400
 19410 19420 19430 19440 19450 19460 19470 19480 19490 19500
 19510 19520 19530 19540 19550 19560 19570 19580 19590 19600
 19610 19620 19630 19640 19650 19660 19670 19680 19690 19700
 19710 19720 19730 19740 19750 19760 19770 19780 19790 19800
 19810 19820 19830 19840 19850 19860 19870 19880 19890 19900
 19910 19920 19930 19940 19950 19960 19970 19980 19990 20000
 20010 20020 20030 20040 20050 20060 20070 20080 20090 20100
 20110 20120 20130 20140 20150 20160 20170 20180 20190 20200
 20210 20220 20230 20240 20250 20260 20270 20280 20290 20300
 20310 20320 20330 20340 20350 20360 20370 20380 20390 20400
 20410 20420 20430 20440 20450 20460 20470 20480 20490 20500
 20510 20520 20530 20540 20550 20560 20570 20580 20590 20600
 20610 20620 20630 20640 20650 20660 20670 20680 20690 20700
 20710 20720 20730 20740 2075

[illegible]

410 420 430 440 450 460 470 480 490
TNIARVLIHRCESLRFDCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----
RIG2A protein VNIARVLIHRCESLRFDCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----DRGRKAI 479
RIG2B protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----PIDDNC 481
RIG2C protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----VM 472
RIG2D protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----DI 470
RIG2E protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----VNRPG 471
RIG2F protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----QAVSLTDE 477
RIG2G protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----VNRPG 488
RIG2H protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----VNRPG 472
RIG2I protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----H 480
RIG2J protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----QAVSLTDE 479
RIG2K protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----GRVQK 478
RIG2L protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----SAFRDYNCHDA 465
RIG2M protein TNIRVLIHTECSLRHFDCCSSIGNLNLEVLSPFANSIGIEMLPSTIGNLKKRLIDLTNCTYGLRIENGVLNVLKLEELYIQNAG-FG-----FKDCHDA 480

[illegible][illegible]

AC15-2A
AC15-2B
AC15-2C
AC15-2D
AC15-2E
AC15-2G
AC15-2H
AC15-2I
AC15-2J
AC15-2L
AC15-2N
AC15-2O

---TGCGAGTTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
410 410 420 430 440 450 460 470 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
460 460 470 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
461 461 470 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
471 471 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
472 472 480 490 500

GTGTCGAGACTGAAATTCCTCTCTCGAGGAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
463 463 470 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
429 429 430 440 450 460 470 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
439 439 440 450 460 470 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
468 468 470 480 490 500

---TGCAGATTCGAAATTCCTATCAGAGAAAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
498 498 500

ATGCGAGACTGAAATTCCTCTCTCGAGGAGAGACATTTGCGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
460 460 470 480 490 500

ATGTACAGACTGAAATTCCTCTCTCGAGGAGAGACATTTACAGAGAGCCCACTGGCA---GTAATTTCTAAT---GTTGTATT---CCCATCTGTCTCATGCA
455 455 460 470 480 490 500

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
510 520 530 540 550 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
559 559 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
557 557 560 570 580 590 600

CTCTTTTCTAATC---CTCGGTGTCTTCAATATGGAATTTATACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
558 558 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
512 512 520 530 540 550 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
560 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
520 520 530 540 550 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
536 536 540 550 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
566 566 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
489 489 500

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
557 557 560 570 580 590 600

CTCTTTTCTAATC---CTCCATAAACTTAAATTCGAGAGATTTGACAGAGTGGAGTGGTGTGTTCTGAGATAGAGAGTGCAGAGTGCAGAGATAGAGAAATTTG
549 549 560 570 580 590 600

AC15-2A	610	620	630	640	650	660	670	680	690	700
AC15-2B	GTACAACTCCATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	653								
AC15-2C	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	651								
AC15-2D	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	652								
AC15-2E	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	654								
AC15-2G	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	654								
AC15-2H	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	614								
AC15-2I	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	630								
AC15-2J	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	660								
AC15-2L	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	580								
AC15-2N	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	651								
AC15-2O	GTACAACTCCGAAATACCAACACACG--CTAT--TATACCTCCCAACCTCCAGAAATGGATCTCTATAGGAATATGGACAACTGAGTCATGTGTGGA	649								
<hr/>										
AC15-2A	710	720	730	740	750	760	770	780	790	800
AC15-2B	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	753								
AC15-2C	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	751								
AC15-2D	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	752								
AC15-2E	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	694								
AC15-2G	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	714								
AC15-2H	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	722								
AC15-2I	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	760								
AC15-2J	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	677								
AC15-2L	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	751								
AC15-2N	AGTGCAGCACTGGATTAATTTCTCTCTTCGAAACACACATCGATATCCCATTTCCACAACTCCACAACTCCATTAATTTATTTATTCGAAAGCAT	737								

SEQ ID NO:

AC15-2A	810	820	TAAGTACTTGTGTTTCACTCTCAAGG - 56	779
AC15-2B			TAAGTACTTGTGTTTCACTCTCAAGG - 57	777
AC15-2C			TAAGTACTTGTGTTTCACTCTCAAGG - 58	777
AC15-2D			TAAGTACTTGTGTTTCACTCTCAAGG - 59	788
AC15-2E			TAAGTACTTGTGTTTCACTCTCAAGG - 60	721
AC15-2F			TAAGTACTTGTGTTTCACTCTCAAGG - 61	781
AC15-2G			TAAGTACTTGTGTTTCACTCTCAAGG - 62	738
AC15-2H			TAAGTACTTGTGTTTCACTCTCAAGG - 63	722
AC15-2I			TAAGTACTTGTGTTTCACTCTCAAGG - 64	784
AC15-2J			TAAGTACTTGTGTTTCACTCTCAAGG - 65	699
AC15-2K			TAAGTACTTGTGTTTCACTCTCAAGG - 66	778
AC15-2L			TAAGTACTTGTGTTTCACTCTCAAGG - 67	763

()

SEQ ID NO: 68

RLG3 (real RLG3)

[Strand]

```
1  AATGGCAAAA GAAGTCGGAG CAAGAGCTAA GTTAGAGCAT CTATTGACG TCATTATCAT GGTAGATGTC
71  ACTCAAGCAC CCAACAAGAA CACAATTCAA AGTAGTATTT CAGAACAGTT GGGATTAAAA CTGCAAGAAG
141 AGAGCTTGTT GSTAAGAGCA GCTAGGGTAA GTGCGAGGTT AAAAATGCTT ACAAGGGTGC TGGTGATATT
211 AGACGATATA TGGTCAAGGC TTGACATGGA GGAACCTGGG ATTCCCTTTG GATCAGATAG ACAACACCAC
281 GGCTGCAAAA TCCTGTTGAC TTCAAGAAGT ATTAGTGCTT GTAACCAGAT GAGAGCTGAT AGAATCTTTA
351 AAATACGAGA AATGCCACTG AATGAAGCAT GGCTTCTTTT CGAAGAACA GCTAAAAAAG CTCCGAATCT
421 GCATCAAGTA GCAAGAGATA TCGTGGAGGA GTGTGGTGGG C
```


RLG4
SEQ ID NO: 69

```
1  GAATTCGGTG  TTGGTAAGAC  AACTCTTGCC  TCTTCTGTTT  ATGATGAAAT  CTCTAGCAAG  TTTGATGGTT
71  GCTGCTTTCT  AAAAAATATCT  GGGAGGAATC  AAGTAATAAA  GACGGTATAG  AAAGATTGCA  AGAAAAAATC
141  ATTTGTGATG  TTTTGAAACA  AGAGCAAGTG  GGGGTAGGGA  GAGTTGAAGA  AGGAAAGCGC  ATGATAAAGG
211  ATAGGTTACA  ACATAGAAAG  GTATTGATTG  TCGTTGATGA  TGTGACAAC  GTTGAGCAGC  TAGCTAGAAC
281  AGTTGGCTGG  ATCACATGAT  TGGTTTGGTG  AAGGTAGCCG  CATAATAATC  ACAACTAGAG  ATGAACATGT
351  ATTAATTGCA  CACAAAGTAG  ATGTGATACA  CAATATAAGC  TTGTTAAACA  ACGATGAAGC  TATGCATCTC
421  TTCTGCAAGC  AAGCACCACG  GGGTCACAAA  CGTATACAAG  ATTATGAGCA  ACTTTTAAAA  CATGTGGTTT
491  CTTATGCTGG  TGGGCTTCCA  CTAGCACTGT  CGAC
```

SEQ. ID NO: 70
FLG1-EL69
[Strand]

```
1 ATCGTAACCG TTGCTACGAG AAGCGCTGTC CTCCTTCATC TTTTGTGATA TGTCATATTC TCATNNATN
71 TGCCACATAT AATTTTGTGG TTATTTTAAA TTAATTTTAA TTCCACATGT CATTTTATGA GTTTTCTAT
141 TTTATTGAGT TTCCATAAAT ATTTAAATGT AATAACAATA AATGCATATT TATTTTCTTT TAAATAAACG
211 CATATAATAT ATAGATTAAA ATCATATAAT ACATAGGTTA AACTCATATA ATACATATGT TCATCCCCAG
281 TTTATTATTA TGCTCATACC TTAATTTATT TATTATTAT TATTAGAGT AGATGATCIT TGTGATATTA
351 AAAATTTAAT TTGTTCAAAA TTTAAATTA TTAATAATCC CACAATTIGA ATAAAAATTA AAAAAATGNN
421 CCCACCATTA GTCCATCACT TTTTCAGCTC ATCAATATCG TGAGTATTCT CCTTCGTTTC CACCTTAATC
491 AATATTTCCT CGGAATGACA GACTCCTACG GCGTTTCCTA ATTTGCGTTC CGACACTGTT CATTTGAAGGA
561 GATAATAAAT CAAATGGAGC TGCTCCAATG TTCAATGCTG ATGAAAGGTG AATTTGTATG GAAGANAATG
631 TCAGCGATCT ATCTCCATCC GGAACCCACC ACATTATCAG TGACCAACA AACCACTCAA AACGGVGGAA
701 GTAGRRACAC WRKAAAGTCA TGAAGAATAG ATTATTTTGT TCCTCATGGG CTGACTGAGG AGCGGGTTTA
771 GTTCATCATT TTTCTTTGAN CAAAGAATTA TCGGTCCATC GAATTTTAC ATCGACAAAG AAGTTTCACT
841 TCGCAATGTT TTGTTAAACA ATTTTAAATC TTTTATCTTT TTGTTGAAA CTCTCAATT GCAACTTGCA
911 ACTTGCAACT TTTGGGCCCC CAAATTTGTG GTGGGCGTTA ATTTAATCCA CATATTCAC GTAAACAATA
981 ATTCAAATCG ATCTCTGTTC ATCCAATPCA TCAACATCTC TTGATAATTG AAATCATCA CGCTTCATCC
1051 ATTTTCATCCA CATCTATACT ATATCTCTCG CTCTTATCAT ATTTAAAGAT GGCTGAAATC GTTCTTCTG
1121 CCTTCTTGAC AGTGGTGTGT GAAAGCTGG CATYTGAAGC CTTGAAGAG ATTTGTTCCGT CCAAAGAAAT
1191 TGAATCTGAG CTTAAGAAAT TGAAGGAGAC ATTAGACCAA ATCCAAGATC TGCTTAAAGA TGCTTCCAG
1261 AAGGAAGTAA CTAATGAAGC CGTTAAAGA TGGCTGAATG ATCTCCAACA TTTGGCTTAT GACATAGAG
1331 ACCTACTTGA TGATTTTGCA ACTGAAGCTG TTCAACGTGA GTTGACCGAG GAGGGTGGAG CCTCTCCAG
1401 TATGGTAAGA AAACATAATCC CAAATTTGTG CACAAGTTTC TCACAAAGTA ATAGGATGCA TGCCAAGTTA
1471 GATGATATTT CCACAGGTTT ACAAGAATCG GTAGAGGCAA AAAATAATCT TGGTTTAAAT GTGATAACAT
1541 ATGAAAGGCC AAAAATGAA AGGTATGAGG CGTCTTTGGT AGATGAAAGC GGTACTGTCT GACGTGAAGA
1611 TGATAAGAAA AAATTGCTGG AGAAGCTGTT GGGGGATAAA GATGAATCAG GGAGTCAAAA CTTAGCATC
1681 GTGCCCATAG TTGGTATGGG TGGAGTTGGT AAAACAATCT TAGCTAGACT TTTGTATGAT GAAAGAAAG
1751 TGAAGGATCA CTTCGAATCT AGGGCTTGGG TTTGTGTTTC TGATGAGTTT AGTGTTCCTA ATATAAGCAG
1821 AGTTATTATAT CAATCTGTGA CTGGGGAAAA GAAGGAGTTT GAAGACTTAA ATCTGCTTCA AGAAGCTCTT
1891 AAAGAGAAAC AAACATAATCC GCTATTTCTA ATAGTTTGGT ATGATGTGTG GTCTGAAAGC TATGCTGATT
1961 GGGAGAAATT AGTGGGCCCA TTCTTGGGG GGTCTCTTGG AAGTAGAATA ATCATGACAA CTCGGAAAGGA
2031 GCAATTTGCT AGAAAGCTGG GCTTTTCTCA TCAAGACCTT CTGAGGGTTC TATCACAAGA TGATGCTTTG
2101 TCTTTGTTTG CTCAACACGC ATTTTGGTGA CCAAACCTTG ATTCACATCC AACACTAAGG CCACATGGAG
2171 AACTCTTTGT GAAGAAATGT GATGGCTTAC CTCTAGCTTT AAGAACAATT GGAAGGTTAT TAAGGACAAA
2241 AACAGAGATG GAACAATGGA AGGAGCTGTT GGATAGTGAG ATATGGAGGT TAGGAAAGAG CGATGAGATT
2311 GTTCCGGCTC TTAGACTAAG CTACAATGAT CTPTCTGCCW CTTTGAAGCT RTTTRTTTGA TAYTGCTCTT
2381 TGTTTCCCAA GGACTATGAG TTTGACAAGG AGGAGTTGAT TCTATTGTGG ATGGCAGAAG GGTTTTGTGA
2451 CCAACCAACT AYAAACAAGT CAAAGCAAGC KTTGGGTCTT GAATATTTTTR AAGAGTTRTT GTCAAGRTCR
2521 TTTTTCACAC ATGCTCCTAA TRRCAATCS TTTTGTGTA TGCGATGACCT AATGAATGAT TTGGCTACAT
2591 TTGTTGCTGG AGAATTTTIT TCAAGGTTAG ACATAGAGAT GAAGAAGGAA TTTAGGATGS AATCTTTTGA
2661 RAAGCACCCT CATATGTCAT TTGTATGTGA GRATTACATA GGTTCACAAA RGTTCGAGCC ATTTAGAGGA
2731 GCTAAAAAAT TGAGAACATT TTTAGCATTG TCTGTGGGG TGGTAGAAGA TTGGAAGATG TTTTACTTAT
2801 CAAACAAGGT CTGTAATGAC WTACTTCARG ATTTACCATT GTTAAGGGTC CTRAKTTTGA TTRRTCTTAY
2871 AATAASYRAG GTACCARAAK TCGTSGGTAG TATGAASCAC TTGCGGTATC TTAATCTTATC WGRAACTTWA
2941 ATCACACATT TACCGGAAWA TRCTGCAAT CTTTATAATT TACARACCTT GATGTNTCT GGCTGTGAMT
3011 ATTTAGTTAA KTTGCCCAAR ACCTTCTCAA ASCTTAAAAA TTTGCASCAT TTTGACATGA GGGTACTCC
3081 KAAKTTRAAK AACATGCCCT TARGGATTGG TGARTTGAAA ARTCTACAAA CTCTCTTMYG TAACATTTGC
3151 ATAGCAATAA CCGAGCTTAA GAACTTGCAM AAYCTCCATG GGAARITTTG TAITTGGCGGG CTGGGAAAAA
3221 TGGAAAAATG HGTGGGATGC ACGTTAAGCG AACTTGTCTC A:AAAAAGGT TWAATGARTT ANAACTGGR
3291 WTKGGGGTGT ATRAAATTTAA TGTTTTCCGA AATGGGAACA CTGAAAAAAA NAAGGTCCTC AATGAATTGA
3361 ATGCCCTACA ATGGTAYTCY AAMWAARRRY YYWTARWAT TWMKAWRRK GKGTTYATRR TKPTMYRAAW
3431 WAGRGTKTXX KARGTAGGTT TCATCCAAATC ACCCAAGTGG GAAAAATAGAT GATATTTTCA GGGCTTACTG
3501 ATGAGATGTT GAGAGGTATG ATAGGNTVIC TTGGGGCGGT AGAAGAAATA AGCATCCATT CTCTTAATGA
3571 AATAAGATAAT YTGTTGGAAAT CAGAAGCAGA GGCAAGTAAG GTTCTTATGA ATTTAAAGAA GTTGGATTTA
3641 GGTGAATGTT AAAATTTGGT GAGTTTAGGG GAGAAAAAGG AGGATAATCA TAATATTAAT AGTGGGAGCA
3711 GCCTAACATC TTTTAGGAGG TTGAATGTAT GGAGATGTAA CAGCTTGGAG CATTCGAGGT GTCCAGATAG
3781 CATGGAGAAAT TTGTATATGC ACATGTGTGA TTCAATNACA TCCGCTCTCT TCCCAACAGG AGGAGGACAG
3851 AAGATCAAGT CACTTACCAT CACTGATTGC AAGAAGCTTT CCGAAGAGGA GTTGGGAGGA CGAGAGAGGA
3921 CAAGAGTGTG TATAAATCA AAAATGCAGA TGCTTGAATC AGTAGATATA CTAATTTGCC CAARTCTGAA
3991 ATCTATCAAT GAATGAGTT GCTTCATPCA CCTGAACAGA TTATATATAT CAAACTGTCC GAGTRTGGAG
4061 TCATTTCCTG ACCATGAGTT GCCAAATCTC ACCTCCTTAA CAGATCGAAG GAGAGGACAG CGATTTTCTT
```

RLGI-E169
[Strand]

4131 ACGAACGGTT ACGATTGAC TGGCGTCGT TTT

SEQ ID NO: 70

Further Characterization of RG2 Family Members:

Further sequencing of cloned RG2 polynucleotide sequences, as discussed above, identified additional RG2 species, listed below. Additionally, further sequencing of the 5' sections of RG2 sequences listed above resulted in modified and/or new sequence information, also listed below. The AC15 sequences found in the 3' sections of RG2 family have not changed.

Listed below are: four full length species, RG2A, RG2B, RG2C and RG2S; two near complete, but with a gap in the largest intron, RG2D and RG2J; three nearly complete RG2 gene sequences, RG2K, RG2N, and RG2O. The deduced translation products (polypeptides) encoded by these RG2 species are listed below. The polynucleotide sequences do not contain any gaps (as with some of the polynucleotide sequences), because all of the gaps in the sequences are in introns, *i.e.*, there are no gaps in exon, or coding, sequences.

They include: an RG2A polynucleotide sequence (SEQ ID NO:87) and its deduced polypeptide sequence (SEQ ID NO:88); an RG2B polynucleotide sequence (SEQ ID NO:89) and its deduced polypeptide sequence (SEQ ID NO:90); an RG2C polynucleotide sequence (SEQ ID NO:91) and its deduced polypeptide sequence (SEQ ID NO:92); an RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94), and its deduced polypeptide sequence (SEQ ID NO:95); an RG2E polynucleotide sequence (SEQ ID NO:96) and its deduced polypeptide sequence (SEQ ID NO:97); an RG2F polynucleotide sequence (SEQ ID NO:98) and its deduced polypeptide sequence (SEQ ID NO:99); an RG2G polynucleotide sequence (SEQ ID NO:100) and its deduced polypeptide sequence (SEQ ID NO:101); an RG2H polynucleotide sequence (SEQ ID NO:102) and its deduced polypeptide sequence (SEQ ID NO:103); an RG2I polynucleotide sequence (SEQ ID NO:104) and its deduced polypeptide sequence (SEQ ID NO:105); an RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107), and its deduced polypeptide sequence (SEQ ID NO:108); an RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110), and its deduced polypeptide sequence (SEQ ID NO:111); an RG2L polynucleotide sequence (SEQ ID NO:112) and its deduced polypeptide sequence (SEQ ID NO:113); an RG2M polynucleotide sequence (SEQ ID NO:114) and its deduced polypeptide sequence (SEQ ID NO:115); an RG2N polynucleotide sequence (SEQ ID NO:116) and its deduced polypeptide sequence (SEQ ID NO:117); an RG2O

polynucleotide sequence (SEQ ID NO:118) and its deduced polypeptide sequence (SEQ ID NO:119); an RG2P polynucleotide sequence (SEQ ID NO:120) and its deduced polypeptide sequence (SEQ ID NO:121); an RG2Q polynucleotide sequence (SEQ ID NO:122) and its deduced polypeptide sequence (SEQ ID NO:123); RG2S polynucleotide
5 sequence (SEQ ID NO:124) and its deduced polypeptide sequence (SEQ ID NO:125); an RG2T polynucleotide sequence (SEQ ID NO:126) and its deduced polypeptide sequence (SEQ ID NO:127); an RG2U polynucleotide sequence (SEQ ID NO:128) and its deduced polypeptide sequence (SEQ ID NO:129); and RG2V polynucleotide sequence (SEQ ID NO:130) and its deduced polypeptide sequence (SEQ ID NO:131); and, an RG2W
10 polynucleotide sequence (SEQ ID NO:132) and its deduced polypeptide sequence (SEQ ID NO:133).

Characterization of New RG Family Groups and RG Species:

Further BAC insert characterization and sequencing, as discussed above,
15 identified new RG polynucleotide sequences. The new sequences were characterized as belonging to new RG families; designated RG5 and RG7. These RG polynucleotides sequences, and their predicted translation products (the polypeptides which are encoded by these sequences) are summarized and listed below.

Identified and listed below is an RG5 family member, designated as the RG5
20 polynucleotide sequence set forth in SEQ ID NO:134, and its deduced polypeptide sequence (SEQ ID NO:135). This sequence contains an NBS region sequence.

Also identified and listed below is an RG7 family member, designated as the
RG7 polynucleotide sequence set forth in SEQ ID NO:136. No deduced polypeptide
sequence is given for the new RG7 family member as this sequence appears to be a
25 pseudogene.

RG2A polynucleotide sequence (SEQ ID NO:87)

AAAGTTCATATCCAAGCTTGCCCTCCAACCTAGCTCCTTCAATGGCACC
TCCTTCTCTTCAAAAGCACACAAGAACACTTTCAAGCTCAACCACACTCA
30 CACAAGCTCTAGAACGAGGGTTAGGGCACATTTAGGGTTTTGCTCTCTGG
AAATGGTGTCTAAAAGTGAGGCCATAATGTTTCCTTATATAAGGCTCACTC
CCACAATTAGGCTTTCAATCTGAACGTANTACGCCAGTGTAACACTATGG
TACGCCCAACGTACTCGGTAGTCTCCGCGTCAANAATACTCATGAGTA

CGCGCAACGTA CTTTCCCTTACGCCCAGCGTACTCAAAGCCAAACATTC
TTTTCAAGGACTAATTTTGACAACCTTGAGGAAAGAAAAGGATCAAAGANA
TATACTTGAATTCCGGGATGTTACAATGAAGTTGANACCTTGGCTAAAAA
ATTAAATTGGTTGTGGAAGCCGTTGGCTGAGCAAGCAACAAGGGTAAAT
5 TCGTAATCTACAAATGGTGTTATTTTCTATTTCTTCTTATTATTTTACTT
GATTTACGGGTAGTTTTTTTTTCTTACAAAAAATATTAAAGTTGATAAAG
TATAGCCACTAAAATTGACTTTTTTCCAAAACATAATGTCAAATGGTGCGT
ATATGTATCATGTTGTATTANATAATGAATATGATGATNCTGTTCTATTT
AANCCGAAAAAATTATCTAATGATTTTATATTGGAACAAAGTTGTGAT
10 TTTTNGCATAATATAATCAAATCCNCTTTTGTNTGGGAGGTGGATAAATG
TGGTAAATTTANAACAAGTGTTTTNACNTTGAAGGGTNTGGAAGGTTGA
AAAAAGTTAAAATGATAAAATGTTTACACAAATGTTGTATCCGACTGAAT
ATNATGTTTAAGGATNATTGTATTAAATTGTTGATATATAGTAAGCATAA
ATATTTAGAATTGTGACTTAAATTTATAAGTTATNCNAACTGGATTGAAA
15 CATTTTTGATATANATTAGGAATGAAAATGAGCAACCCTAACATACTTAT
CTTTGGTAGTTTGGTTATTATATTTTTATTANAATATAGAANCATCCCTT
TATTTTAAACCCATATTGTGGACGGACTTGAATAAATGGGAAAAATGTAC
CTTGCTATTTAGCACAAAAAAATTATAAAAATGTACATTGCTATTTAGCA
CAAACAAAAAATAAATACTTATCCTTTTTGCATTAGGTCACAAAGAAATA
20 TAAATGGGAAATGTGTTGCTATTTAATGCACTAAAAGAACTATTTTGC
CTTTATTAAACCGGGTAAACCAATAGAAAAATGGAAGTACATTGTCATTT
AGC.ATGAAAAAAATAA CTTTCCATTTTTTGCATCCGGTCACAATAATAG
AAAAATGAAAGTACGTTGCTATTTAGCGAACTAACTTCCTTTTTTCTTT
TTGGCATCGTATCATAAAATATAGACTAAAATACGTTAGTTTTACATTTT
25 TAATACATTGAAATGTCTAATCCACATGTTATTCTATAAAAAGGGAAATG
TAATTTACTTATTCTTTGATTCTTTGGCTTCTTTTTAGTACCCAAAACAT
CCCTCTATCCATCTATTCCAATAAATAATGAAAATATATTCCTTCCA
TTGTAGGGATGTTATAAATTTTGTAAATTGTTTTTATGCAAAAAAGTGTTT
TTTGTTAACTAGATTAACGAGATTCATTTTTCAGCATTTTAGGAGAAGTT
30 CATCCATCTTTTGGATATGAAGTGCAAGCCAAGTTCTTTAACATGGAATA
TGAGGTCCCTATATGCTCAAAAAATAGCAAATGAGAAATTTTTTAAATTG
GATCCCCATAAAAGAAAATTTGTAAATGGTTGTTTAAATATTGGTCAATG
TGTCCACCGGATGAGCATAATACTAGTTTATAAGGGGTAAAGGTGGGTTT
GGTGGGCCCATTTATCTTTATTATTTCTAAAAGTCAGAATTAAGTAAAAA
35 AAATTATAAGATAAATAACCATAAGGATAAAAAATCATTTTATTTGGACCA
AAGACCAAAGTTGTAAAGGGGCTGTTTGTTTTTTTTGTGAAGAGCTGTGC
AACCACTTTTGTCTGCGCCGCACAGACAACGTGCAGACATATGCCCTCGC
AGAGTGTTTGTTTTTTGAAGTGCGCAGACCAAAAAAACGTCTGCGCGAG
GTCATCCTGGCGCATATATGTGTCACTGTCTTCAAAGGTCTTCAGACCTC
40 ATTTTAACCAAAAAAAGACCACCGGTTTTTTTTTTTTTTTTNTTC
TTTCTCTGTAGCTGAAAATGCATTTTAAATCTTTATGACATGAAATTAA
GTTTGAAAAATTAATTTATTTCAACAGCTGTAGACGTTAAAAACAAACAG
TCTTCTTGTTGCAGACTGTGGACATTTGGTCCACCTCTTCTACCGCAGAG

ACTTGCAGATGTGGTCCGCAGACTGCAGACATTTTGGCTTCAAATAAACA
AACATCACCTAATTTGACTACACCACACGGACCTCCAATGTAACAAAAA
AAGGTTGAAACAAAGTTGCCTATTTCTCCATATCCAGGGGCCATTTATGT
AAGAGTTATCTAAATTTTAGTTTCGGTAGATCAGTTCTCACATTTTAACCG
5 GGTAAAGTGTATGTGTGTACGCGCGCACCTGAAAGGTTTGAANGTAACTT
CCAACTGAANCAANAATCGATATGAAGTATCAAGTTAGAGGTTCAATTG
GTG.AAGGAATCAGCTGGAGGTTGGGGAATCGAGCTTCCACTATTAAGGTA
AAATCCATAACCCTAAATGTTGGTACGCTCATATATCAAATTGCGTGTTT
TGTTGAATGAAAAAAGCATGCTCAAAAAACCAGTGTAAGGCACGGTATAT
10 GACATATTTATAGTTACTGATAACAAATTATGATAATTTTGGGTTTACGT
AAGTTAGGATTTCGTACTTCAACCAAATGTAATAGTTTTTGTGAGTCTATC
TATGTATTTGGGGAATCACATTAGCAACGGGATTGTACTAGTAATTCGAA
AAAGTCTTTTAAATAATTTTCTGTTTATAATTTATGAATAGTTTTAGCG
ACATCTAATATTAATAAGAAATGTATCTGATATTGAATTAATGTCCTTAAT
15 GTG.AACATAGACCTTTTCCATTTACTAATGCCTAATTATTAGTTTCTAAT
CAATAAATTTTAAATTTCTGTTTTATGCTTCTAAGACAATAAAAAATCCATG
ATTTACCTTTAAATATTAACAAAAATGACCATAAATAAAAAAATTAG
GATACCAAACCCCCCGCCATGCCCAATGTCTAAATATTCTTGATGCTTT
TGCTTTTCCCTCTTTTCCCTTGTTAGTCTATTATTCTGGAGAGTTTGAGAG
20 AGTTTCATACAAGAAAATTTCAAGAAGAAAGCAAAGGTCCAGGTATTCTC
TTTTCTTAATTATGTATTAACCTTACAAGCATTTTTTACACGATCCATGGT
TTTTTGTGTATGTTTTTCAAATTGAAACTAGATTGGGACTTTTGCCCTTG
ATG.ATTCATAAGATATTGCATGGAGTTGAGATTGTGTAAGAAAAGTGGTG
AATAGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATATGATGAT
25 GAGATAGAGATATGTTAAAACCTGGCTAGAAAATTGTTTTAATTTGAAATT
TAGGTTGTTGAATTTGAAAGATACCAAGCTAATAACTAATTAGTTATGCT
AAATAGTTATAAAGAACAACAACTCGTAGTTTTTTTTTCATGATTTTCA
ACCTCTTCGTACCAAACCTAAATTATAACAAAATTGAATATCATTCTCTGC
AATCAATTTTAACTTTTGTTATTATCATCATGTCTAAAATTGCCACAAGT
30 TTATTTTCATAGTCATATTGGATTATGAAAGGACTATTTTACCAATTAC
ATCTTTACTTTATGGCCAAAGCTAATAACAATCCGACTAAACTAAAGGATT
CTAGGATGCATATAGTTTGCTCCCCGATTATAGATTTCTATCTAATTTGT
CTATTGTACTAATTTAGGTGCCACCACAAGTAAATTCCTGAAATGGATGT
CGTTAATGCCATTCTTAAACCAGTTGTGCGAGACTCTCATGGTACCCGTTA
35 AGAAACACATAGGGTACCTCATTTCCCTGCAGGCAATATATGAGGGAAATG
GGTATCAAAATGAGGGGATTGAATGCTACAAGACTTGGTGTGCAAGAGCA
CGTGAACCGGAACATAAGCAACCAGCTTGAGGTTCCAGCCCAAGTCAGGG
GTTGGTTTGAAGAAGTAGGAAAGATCAATGCAAAAGTGGAATTTCCCT
AGCGATGTTGGCAGTTGTTTCAATCTTAAGGTTAGACACGGGGTCGGAAA
40 GAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTGATGAGAGAACT
CTATCATCATTTGGAATGATCATTCCATTCTTTAGGAAGAATTGATTCC
ACG.AAAGCATCCACCTCAATACCATCAACCGATCATCATGATGAGTTCCA
GTC.AAGAGAGCAAACCTTTCACAGAAGCACTAAACGCACTCGATCCTAACC

ACAAATCCCACATGATAGCCTTATGGGGAATGGGCGGAGTGGGGAAGACG
ACAATGATGCATCGGCTCAAAAAGGTTGTGAAAGAAAAGAAAATGTTTAA
TTTTATAATTGAGGCGGTTGTAGGGGAAAAAACAGACCCCATTTGCTATTC
AATCAGCTGTAGCAGATTACCTAGGTATAGAGCTCAATGAAAAAACTAAA
5 CCAGCAAGAACTGAGAAGCTTCGGAAATGGTTTGTGGACAATTCTGGTGG
TAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTTGTGGATCTGA
ATGATATTGGTTTAAGTCCTTTACCAAATCAAGGTGTCGACTTCAAGGTG
TTGTTGACATCACGAGACAAAGATGTTTGCCTGAGATGGGAGCTGAAGT
TAATTCAACTTTTAATGTGAAAATGTTAATAGAAACAGAAGCACAAAGTT
10 TATTCCACCAATTTATAGAAATTTCCGGATGATGTTGATCCTGAGCTCCAT
AATATAGGAGTGAATATTGTAAGGAAGTGTGGGGGTCTACCCATTGCCAT
AAAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCATGGAAGA
ATGCACTTCTTCGTTTAGAGCACTATGACATTGAAAATATTGTTAATGGA
GTTTTTAAAATGAGTTACGACAATCTCCAAGATGAGGAGACTAAATCCAC
15 CTTTTTGCTTTGTGGAATGTATCCCGAAGACTTTGATATTCTTACCGAGG
AGTTGGTGAGGTATGGATGGGGGTGAAATTATTTAAAAAAGTGTATACT
ATAGGAGAAGCAAGAACCAGGCTCAACACATGCATTGAGCGGCTCATTCA
TACAAATTTGTTGATGGAAGTTGATGATGTTAGGTGCATCAAGATGCATG
ATCTTGTTTCGTGCTTTTGTGTTTGGATATGTATTCTAAAGTCGAGCATGCT
20 TCCATTGTCAACCATAGTAATACACTAGAGTGGCATGCAGATAATATGCA
CGACTCTTGTAAGAACTTTTCAATTAACATGCAAGGGTATGTCTAAGTTTC
CTACAGACCTGAAGTTTCCAAACCTCTCCATTTTGAACTTATGCATGAA
GATATATCATTGAGGTTTCCCAAAAACCTTTATGAAGAAATGGAGAAGCT
TGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCCTCATCAC
25 CTCAATGTTCCGTCAACCTTCGCGTGTTTCATCTACATAAATGCTCGTTA
GTGATGTTTGACTGCTCTTGATTGGAAATCTGTGCAATCTAGAAGTGCT
TAGCTTTGCTGATTCTGCCATTGACCGGTTGCCTTCCACAATCGGAAAGT
TGAAGAAGCTAAGGCTACTGGATTTGACGAATTGTTATGGTGTTCGTATA
GAT.AATGGTGTCTTAAAAAAATTGGTCAAACCTGGAGGAGCTCTATATGAC
30 AGTGGTTGATCGAGGTCGAAAGGCGATTAGCCTCACAGATGATAACTGCA
AGGAGATGGCAGAGCGTTCAAAAGATATTTATGCATTAGAACTTGAGTTC
TTTGAAAACGATGCTCAACCAAAGAATATGTCATTTGAGAAGCTACAACG
ATTCCAGATCTCAGTGGGGCGCTATTTATATGGAGATTCCATAAAGAGTA
GGCACTCGTATGAAAACACATTGAAGTTGGTTCTTGAAAAAGGTGAATTA
35 TTGGAAGCTCGAATGAACGAGTTGTTTAAGAAAACAGAGGTGTTATGTTT
AAGTGTGGGAGATATGAATGATCTTGAAGATATTGAGGTAAAGTCATCCT
CACAACTTCTTCAATCTTCTTCGTTCAACAATTTAAGAGTCCTTGTCGTT
TCAAAGTGTGCAGAGTTGAAACACTTCTTCACACCTGGTGTGCAAACAC
TTT.AAAAAAGCTTGAGCATCTTGAAGTTTACAAATGTGATAATATGGAAG
40 AACTCATACGTAGCAGGGGTAGTGAAGAAGAGACGATTACATTCCCCAAG
CTGAAGTTTTTATCTTTGTGTGGGCTACCAAAGCTATCGGGTTTGTGCGA
TAATGTCAAAATAATTGAGCTACCACAACCTCATGGAGTTGGAACCTGACG
ACATTCCAGGTTTCACAAGCATATATCCCATGAAAAAGTTTGAAACATT

AGTTTGTGAAGGAAGAGGTAATATAAATTTTAAATGCTAATACATTAC
AAAGGATCTTTTCAGTTAAATCTTTCAAAATATATTGTAATTTGATTGTA
TGGGGTATTATTGTTGGATGGGACTATTAATAAATGATTATCTTGCAGGT
TCTGATTCTTAAGTTAGAGAAACTGCATGTTAGTAGTATGTGGAATCTGA
5 AGGAGATATGGCCTTGCGAATTTAATATGAGTGAGGAAGTTAAGTTCAGA
GAGATTAAAGTGAGTAACTGTGATAAGCTTGTGAATTTGTTTCCGCACAA
GCCCATATCTCTGCTGCATCATCTTGAAGAGCTTAAAGTCAAGAATTGTG
GTTCCATTGAATCGTTATTCAACATCCATTTGGATTGTGTTGGTGCAACT
GGAGATGAATACAACAACAGTGGTGTAAGAATTATTAAAGTGATCAGTTG
10 TGATAAGCTTGTGAATCTCTTTCCACACAATCCCATGTCTATACTGCATC
ATCTTGAAGAGCTTGAAGTCGAGAATTGTGGTTCCATTGAATCGTTATTC
AACATTGACTTGGATTGTGCTGGTGCAATTGGGCAAGAAGACAACAGCAT
CAGCTTAAGAAACATCAAAGTGGAGAATTTAGGGAAGCTAAGAGAGGGTGT
GGAGGATAAAAGGTGGAGATAACTCTCGTCCCCTTGTTTCATGGCTTTCAA
15 TCTGTTGAAAGCATAAGGGTTACAAAATGTAAGAAGTTTAGAAATGTATT
CACACCTACCACCACAAATTTTAATCTGGGGGCACCTTTTGGAGATTTCAA
TAGATGACTGCGGAGAAAACAGGGGAAATGACGAATCGGAAGAGAGTAGC
CATGAGCAAGAGCAGGTAAGGATTTCAATTTCACTGTCTTAATTAATGAT
TAAGCTCCTGCTTTTTGAATAAAAAAGGGACAAACCATTTTCATGACTTAA
20 TGTAGCAATACAAGTCATGTATAAGAGTGACCAACTCTTTTTTATTTATA
AAATGACTACAAAATATTTTTTTTCATTAGAGATCATGTATAAATGTGAC
TAATTTTTTCATCACCTAACTTTAGTTGATAAATCTTTATAAATGTCTACTA
GTTACTTTTCAGTAAAATAACAAATTTAATAAATTATCAACAAAAAGCAT
CAACTAAAAAATCCCACAACCCGTAATAATTTAAAAATAAAAGGATTTAA
25 CATCTAATACGAACAATTTTTTTTCTAAACATGATTTGGACCAAATATCA
CCAGCAACTCAAGTTTGAATCGATTTCAGCTTAAAACTTGACCAGCATAA
TTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATATAAGTTCGTTTCATC
TTTTTTCTTGATCTTGATAGCAAGTTGAATGATTTTCTTCTTCAAAATTG
ATAAAAATCTACATTATAAAGAGACTAGCTTGAAAAAAATGGTCTAGGT
30 GGGTCTTGGGTCTGCTAGATGAAGATGGAAGGGGAGAGTAGATTTCAAA
GACACAACACATCCTTCATTTTATTTATTATTATTATTATTTTTTG
ATATCTTGCTCATATTTGTTACAGATATGTGAGGTCTATTAATCTTTTTA
AATATATAAAAAAATAAATAACATAAATGAGAAAATTAAATAAAGAATAA
ATTAATAAGGGCACAATAGTCTTTTTTAGGTAAGACAAGGACCAAACACGC
35 AACAAAAATAAACAGTAGGGACCATCCGATTTAAAAAAAATAATTAGGGA
CCAAAAACATAAATTCCCCCAAACCATAGGGACCATTTCATGTAATTTACT
CTTACTTTTCGTTTTGTTTCATATTTGGGTAACTATTTTTTTGTACACAT
CTAGGTAACGAACTTGTTGAAGTGTTCCCATTTAGGATGTGACCTACTAC
AACCGATCATAATAGTCATATGTGAACACTTCCAACAACCTTTATTACTTA
40 GGTGTGTACAAAAAACAATAGTTACCATGATGTGAACATACTGAAAAAT
TAATTACCTTAGCAAGTTATTTTCCCATTTAGGTTGTATGGAAACAGTTC
CGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAACTTAACCCTTC
AATTAACCTACCTTTTTTCTTATTAACCTCAATTTCAACCTAAATTCTGATT

CTTGTTTGAAAGTAAGTTGCATCTTTATTTTTGTATTATCTTGTTGCATA
GGATCCTTAGCATCTTTAATAATTTATTTGAAGGTGAAAGATCCAACTA
TTTTTAATCTGTTGGCATTTCATCATTTGCAACTGTTTCTTGAAAAAA
AAATACCTAAAAATCAAAATAACCATTTTCAAATCCAAAATTATAAGAGAG
5 AATTGTAAATGGACATGGAATCATAAATCATTAACACAGTTCAGTAAACA
AGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAAA
GAGACATTACAAGAAGCCACTGACAGTATTTCTAATGTTGTATTCCCATC
CTGTCTCATGCACTCTTTTCATAACCTCCAGAACTTATATTGAACAGAG
TTAAAGGAGTGGAGGTGGTGTGTTGAGATAGAGAGTGAGAGTCCAACAAGT
10 AGAGAATTGGTAACAACCTACCATAACCAACAACAACCTATTATACTTCC
CAACCTCCAGGAATTGATTCTATGGAATATGGACAACATGAGTCATGTGT
GGAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCA
GAATCCCCATTCCACAACCTCACAACCATAAAAAATTATGTATTGCAAAAG
CATTAAGTACTTGTTTTCGCCTCTCATGGCAGAACTTCTTTCCAACCTAA
15 AGCATATCAAGATAAGAGAGTGTGATGGTATTGGAGAAGTTGTTTCAAAC
AGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACACAAC
CACCCTTTGTTCCCTAGTCTTGATTCTCTCACTCTAAGTTTCCTGGAGA
ATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAA
ATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCAATTTGA
20 GGTATGCTTTGTACATATTCAATTATTTATTTAATTTCTTTTTTTATTG
CAATATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAAT
TACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAACTCT
AGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGAT
GGGTAATATAGGCAATTTAAGTTTATTTCTGTAAAGCAGTATTTAGCA
25 AGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGTC
ATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATCAGGGGTCAT
CAGGTGACAGATATTGTAGAATAGAACAATATATAATATCACCCAAAAC
ATTTTTTCTAAGGTTATTCTGTAAATATGTGCTTTCTTGTTTTTCATNGA
ATTNGCATTCGTATATTTTAGGTGTTAAAGTGATTTTNTCTTCAATAAAT
30 CCCGAAATTAATTAACAAAAACAAAAGTACATTTTGTATGTGGAG
AGCACTGGTATCACTTAGTATATAAAAAGCTTGATTTTGAATTAACCTTC
TTATACAAAAGTTGTGTATATAGTTTAATTAGTTTACATCATTTTCCA
TGTGGTGTGTCAGTTGTCTGAAGCAGGTGGTGTCTTGGAGCTTATGCC
AATACGCTAGAGAGATGAGAATAGAATTCTGCAATGCATTGTCAAGTGTA
35 ATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGACAGT
AAGTGATTGCAAAGGGATGAAGGAGGTATTTGAACTCAATTAAGGAGGA
GCAGCAACAAAAACAACAAGAGTGGTGCAGGTGAGGAAGGAATTCCAAGA
GTAATAACAATGTTATTATGCTTTCTGGTCTGAAGATATTGGAAATCAG
CTTTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGAAAGCC
40 TGAGACAGCTCCAAGAGTTAAAGATAACATTTTGCTACGGAATGAAAGTG
ATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGTAAACAACAACAAC
AACACAATAACGAAGGGGGCATCATCATCATCTTCTTCATCTTCTA
AGGAGGTTGTGGTCTTTCCTCGTCTCAAATCCATTGAACTAAATGATGTA

CCAGAGCTGGTAGGATTCTTCTTGGGGAAGAATGAGTTCCGGTTGCCTTC
ATTGGAAGAAGTTACCATCAAGTATTGCTCAAAAATGATGGTGTTCAG
CTGGTGGGTCCACAGCTCCCCAACTCAAGTATATACACACAGAATTAGGC
AGACATGCTCTTGATCAAGAATCTGGCCTTAACCTTCATCAGGTATATAT
5 ATTTCTTTAATTGGCATCATCTAATTAAGAAAGATATCATTCTGCCAAG
TAAATTTACTTCAAACACATTCACACTGGTTTCAGTCTAAGTTTATGTTG
TTCTAGGAAGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTT
CAGTGGAAAGGGTATTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGG
CTTTTATAGTACCTGGAATCGTGTGTGGGAGGAGCATTATTATTCTGATTT
10 GCTTGTTTCTTTATCATTTTTTCTTAGCCTCTGGAACAGCTAGAAACCCT
TTTAACTCTTTTGATTTTCAATGACAAAATTTTCTGTTACTACATTTGA
TTGTTGTTCTTCATGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCT
TTTGATTGTTATTTTCATATCATGTTAGTCACCTGAATCAAGCTTTTCTA
TTTTCAACCAGGGCAAAGGTCAAAGTAACCTACTTTATGAGATCAAAA
15 ACAGCAACCCATCGGATAACTTTTAGTTGGAGTTAATAGTTACAATTACC
ATTGTGATTAATAATTATAATATCTTGTATTAATTCATAAAAATTGGTAC
AGCACATATATGACATTTCAAAGGTTTTTGTGTTGACATATATATGCCTCT
GGCGTTTTCTTTATTGGACATGCAGACCTCATTCCAAAGTTTATACGGTG
ACACCTTGGGCCCTGTAACCTCAGAAGGGACAACCTGTTCTTTTCATAAC
20 TTGATCGAATTATATATGGAATTTAATGATGCTGTAAAAAGATTATTCC
ATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAGATTCATGTGACTT
ATTGTAATTGGGTAGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGG
AGAAATGGAAATAGTGAATTGGTTTTGATGAATCGTCACAAACAACCTAC
CACTACTCTTGTCAATCTTCCAAACCTCAGAGAAAATGAAGTTATGGTATC
25 TAAATTGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTGAG
TTTCCAAACCTAACAAGAGTCGATATATGGGGATGTGATAGGTTAGAACA
TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAC
GCATATGGAAGTGCAGTCAGATAGAGGTCGTGATTGTTTCAGGATGCAGAT
GTTTGTGTAGAAGAAGACAAAGAGAAAGAATCTGATGGCAAGACGAATAA
30 GGAGATACTTGTGTTACCTCGTCTAAAGTCCTTGATATTAACACCTTC
CAWGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTA
TTGGATACYTTGGAAATCTACRAATGCCCAGCAATAACCACCTTCACCAA
GGGAAATTCRCCTACTCCACAGCTAAAAGAAATTGAAACAMATTTTGGCT
TCTTTTATGCTGCAGGGGAAAAAGACATCAACTCCTCTATTATAAAGATC
35 AAACAACAGGTAAACCAGATCTTTGTTGCTTNNATAATTCTTAAACNACA
TNTGAAAAGCTTCATGCAAGTTTTTTTTNGTTATATNGTCAAAAACCGCAA
CCTACATTTTCAGCTTTANATTTATGTACTTTATGCAGGATTTCAAACAA
GACTCTGATTAATGTGAAGTGAATATTAAGGTAAATTATATTTTCATGT
TCCTAGTNGCCTATTAATTAAGGCCTTTTAGTTTCGNGATTTTTGGATGT
40 ATTCCTTCATGATGATGTCAATCTTCTAATACCCCATTCATTGTTTGGTTG
AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTTCATCA
TATGAAGGACATTAAAGAACATGGATGCTCTGAAGATGTTGGGAACACA

RG2A deduced polypeptide sequence (SEQ ID NO:88)

MDVVNAILKPVVETLMVPVKKHIGYLISCRQYMREMGIKMRGLNATRLGVEEHVN
RNISNQLEVPAQVRGWFEVVGKINAKVENFSPDVGSCFNLKVRHGVGKRASKIIEDI
DSVMREHSIIWNDHSIPLGRIDSTKASTSIPSTDHHDEFQSREQTFTEALNALDPNHK
5 SHMLALWGMGGVGKTTMMHRLKKVVKEKKMFNFIEAVVGEKTDPIAIQSAVADY
LGIELNEKTKPARTEKLRKWFVDNSGGKKILVILDDVWQFVDLNDIGLSPLPNQGV
DFKVLLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIG
VNIVRKCGGLPIAIKTMACTLRGKSKDAWKNALLRLEHYDIENIVNGVFKMSYDNL
QDEETKSTFLLCGMYPEDFDILTEELVRYGWGLKLFKKVYTIGEARTLNTCIERLI
10 HTNLLMEVDDVRCIKMHDLVRAFVLDMYSKVEHASIVNHSNTLEWHADNMHDS
KRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFKPNFYEEMEKLEVISYDKMKY
PLLPSSPQCSVNLRVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDRLPSTIGKLG
KLRLDLTNCYGVRIIDNGVLKKLVKLEELYMTVVDRGRKAISLTDDNCKEMAERS
KDIYALELEFFENDAQPKNMSFEKLQRFQISVGRYLYGDSIKSRHSYENTLKLVLK
15 GELLEARMNELFKKTEVLCLSVGDMNDLEDIEVKSSSQLQSSSFNNLRVLVSKC
AELKHFFTPGVANTLKKLEHLEVYKCDNMEELIRSRGSEEETITFPKLFSLCGLP
KLSGLCDNVKHIPLQLELELDDIPGFTSIYPMKKFETFSLLKEEVLPKLEKLHVSS
MWNLKEIWPCEFNMSSEEVKFREIKVSNCDKLVNLFPHKPISLLHHEELKVKNCGSI
ESLFNIHLDCVGATGDEYNNSGVRIKVISCDKLVNLFPHNPMILHHLEEELEVEN
20 GSIESLFNIDLDCAGAIGQEDNSISLRNIKVENLGKLREVWRIKGGDNRPLVHGFQS
VESIRVTCKCKFRNVFTPTTTNFNLGALLEISIDDCGENRGNDSEESSHEQEIEILS
EKETLQEATDSISNVVFPSCLMHSFHNLQKLILNRVKGEVVFEIESESPTSRELVT
HHNQQQPHLPNLQELILWNMDNMSHVWKCNSWNKFFTLPKQQSESPFHNLTIKI
MYCKSIKYLFSPLMAELLSNLKHIKIRECDGIGEVVSNRDEDEEMTTFTSTHTTTT
25 LFPSLDSLTLFLENLKCIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSW
SLCQYAREMRIEFCNALSSVIPCYAAGQMQLQVLTVSDCKGMKEVFETQLRRSSN
KNNKSGAGEEGIPRVNNNVIMLSGLKILEISFCGGLEHIFTFSALESRLQQLKITFC
YGMKVIVKKEEDEYGEQ.TTTTTITKGASSSSSSSSKEVVVFPRLKSIELNDVPELV
GFFLGKNEFRLPSLEEVTIKYCSKMMVFAAGGSTAPQLKYIHTELGRHALDQESGL
30 NFHQTSFQSLYGDTLGPVTSEGTTCSFHNLIELYMEFNDAVKKIIPSELLQLQKLEK
IHVTYCNWVEEVFETALEAAGRNGNSGIGFDESSQTTTTTLVNLPLNREMKLWYL
NCLRYIWKSNQWTAFEFPNLTRVDIWGCDRLEHVFTSSMVGSLLQLQELRIWNCSQ
IEVVTVQDADVCVEEDKEKESDGKTNKEILVLPRLKSLILKHLPCCLKGFSLGKEDFSF
PLLDLTLEIYKCPAITTFTKGNSTTPQLKEIETHFGFFYAAGEKDINSSIIKIKQQDFKQ
35 DSD.CEVNIK

RG2B polynucleotide sequence (SEQ ID NO:89)

TTTTTTAAGATCAGGGATTCAAATTCAGCCCTAGTGATTACAATTGTGTC
TAAACTTTCCCATACCTTCACATTATTGTAAGTATACTTTCTCAGTTTCT
40 CTCTTGGAAGCTTCCTTGGTATTTTAACTCGTGTTCTAATATTTAACTCT
GATAGTTATTTTGGCCAATCTACTATCTGCATGTCCGGTTATTGAATCCG
AAGGCACTGGAATCTTGGATTCCATTCCGTTGTGTGTTTGGTTGCCAAAT

GAACGGAATTGAATTATGTAAGATTCCTTCAAAAATCCATGTTTAGGTATA
TCGTTGTTTCTTGGGATGGATGGTAAAGAACGGAATTTCTCCTGTTCAATT
TTTTAATGAAAGACCAAATTGACCTTATAAACCTGTAAAAAAATTACAT
TCCAGTTTTCTTAACAACTGAAAATGGTAAAGGAGTGTGATTGAATTCC
5 AATCTGTTTCCTGTCCAAAACACGTGACGGAATATTACAATTCCTTCAAA
TTTCATTTTCTTAAATTGTTATTCCCTTTCTTACAAAAACAAGGTAAACG
AAACACCCCGCTTACTTAATCATACTCCTACATGATGTAAATGAAAAGGGT
ATAAATGGTATTTTATTACAGGGATGAGTCACCATGGTCATGAAAGAAT
CATTAACCGCCCTTACCCAATTCATGTTTGCCCTAAAATATGATTTAAA
10 GTAATATTGGCTTATGGGATTCAAGTTGACTTTTTTGTGGCGAAGAAATA
ATGAAAATCTTCATTTCTAAAGTGTCTTCTACCACTGACATTTTCTAAGA
AAGAACTTGCTAGAAGAAGGTGGGTGTTTAGTCTTTTACTCTTTAAAT
GTGAAGACTGTTGAGTTATTATTATTATTTGCCAACTATGGACAACCTG
TTTAGTTTTTTTTTTTCCCAATATCCATTTATATGCGATTTATTTCTGA
15 AATAATTTTATCAAAACGCAGGAAACAATGTAGAATAATACTGGTATAAT
TAATTATATAAAGTTATTAGGCTGAAATCTTGAGGCTACTATAATTTAAT
TATCATAATTTGAAAATCATCAAATTGTATTCCATGTATATTTATGTTAT
CAGATAATTAATAATATGTGAGCCACACAAATCCACATCATCAGACACCC
CACCTTATTGTGCGCTACCTCACCACTTGCATGATCCCGACATCTTCCCA
20 ACCCCACCGACGACTTGGGGTCTCCTTAATATATCAATTATTTTCTGTAA
GTATTTATTTGTGTAAATGTGTAATGTCATTTTACCTTTTTTCTAATATA
TACAGAAACATAAATTTTAAATGAAATTCAACTGCGTTTCATTCTTGCAT
TAAAAAAAAGACTGTACTGTTGTCAATATTTTACTTATAACCTGATTAA
TTAATTAAAGCGTAATTGCATAATTTGCATTAGGTTGTAATTTTGTGTTT
25 TAT.AGGGAGGGTGAGGGTCACCGGAATCAAAGCACTTATGTAAAAGCAG
GGA.AATACAAAAAATTTACTCGAAACAAATTTTATTCAATTTAAGTGAGA
TAATAATGTTCTGATTAGATTATGAGAACTAGGAGATTTAAGTGATATAT
CCC.ATTTAAAAGAAATTGCATTATTAATTTTGGATCTCTTGATGATGACA
AAATTAACCTCGTGACAGGTTATATATCATATACAAAATGAGTGGCTATGC
30 TTTCGCTTTCCAAAAAGCAATTATAGTTATACTACACCTACAAATTTTAA
AAGGGGTAAACATATCAAAATACTTGATAAGTAATTATATAAATATGCA
TTT.AACCCTCTAAAGAAAATGCTACTAAGCTTGGACCATCTCAGAATTAC
AATCATACCCTTCCCCTCAAAAAAGATTTCGTATATATCATGTCATTTGGC
ATTCATTTCTTTTTTCACAATTCATAGTTCTATTCTCAAAAAATTCGAGTT
35 CTCGTATTTGTAAGGAAGATCAGAAGAGACTGTTACACAGGTACTCTCT
TTT.ATTTATTGATTCACATTCATATATGTTATTGTTTTCTTGCTTAATGG
TTTCGTCACTCTAACTGCGCTTGCTGATTAAATTTCTTCACTTTCTTCC
ACGGATTTTTTAAATATTAGTTTTGTGAATGAACAATTGGTGAAGGAAAG
AAACATGGGAGTCTTTTCTAAAGTAAACCTAGATACTTAGGTTATAAGGG
40 TAT.ATGCTAAATGAACTATGCCCATTCACCTTTGCCTTTTCTTTTACTT
TTTAGTTTTTAGAATCCAAGTTTTCATATGTATCTCGATGTGTGAGAAGA
ATAGGCATTAGAAAGGTAAAGGACGTACATAAAATTGATTAATTAGTGAA
TGTTCTTTGATATCATTATTTTTACTCTCATAAAAAGCATATAGATCAAA

CACAAATTGCTACTTGTTAGTGTAACAACCTTCGACTTAATAATGTTAATA
ATCAAGATTCTCTTGATTTCAACTATTTTCTAACCGAACAAGCTCACTAA
AAACTCATATTGCTTTGAGTCTGAGTGGTTTATATTTGGGGTTTTACATT
TAAATTTTTGTGCATGAATGTGAAAATAGACTGCTTATTGATTCTTTGTG
5 TTTCAATTGAGTTGATTTTCATTATTACTACCTTACAAATTGCTCAGTGAT
AGATTTCCATTAATTTGCTAATTCGGTTGCTTCTAAATATGTAGGAGCTA
CTAAAAGCAAAAATATCGAGCAATGTCCGACCCAACGGGGATTGCTGGTG
CCATTATTAACCCAATTGCTCAGACGGCCTTGGTTCCCGTTACGGACCAT
GTAGGCTACATGATTTCCCTGCAGAAAATATGTGAGGGTCATGCAGATGAA
10 AATGACAGAGTTGAATACCTCAAGAATCAGTGTAGAGGAACACATTAGCC
GGAACACAAGAAATCATCTTCAGATTCCATCTCAAATAAGGAATGGTTG
GACCAAGTAGAAGGGATCAGAGCAAATGTGGAAAACCTTCCGATTGATGT
CATCACTTGTTGTAGTCTCAGGATCAGGCACAAGCTTGGACAGAAAGCCT
TCAAGATAACTGAGCAGATTGAAAGTCTAACGAGACAACTCTCCCTGATC
15 AGTTGGACTGATGATCCAGTTCCTCTAGGAAGAGTTGGTTCCATGAATGC
ATCCACCTCTGCATCATTAAAGTGATGATTTCCCATCAAGAGAGAAAACCTT
TTACACAAGCACTAAAAGCACTCGAACCCAACCAAAAATTCCACATGGTA
GCCTTGTGTGGGATGGGTGGAGTGGGGAAGACTAGAATGATGCAAAGGCT
GAAGAAGGCTGCTGAAGAAAAGAAATTGTTTAATTATATTGTTGGGGCAG
20 TTATAGGGGAAAAGACGGACCCCTTTGCCATTCAAGAAGCTATAGCAGAT
TACCTCGGTATACAACCTCAATGAAAAAACTAAGCCAGCAAGAGCTGATAA
GCTTCGTGAATGGTTCAAAAAGAATTCAAGATGGAGGTAAGACTAAGTTCC
TCATAGTACTTGACGATGTTTGGCAATTAGTTGATCTTGAAGATATTGGG
TTAAGTCCTTTTCCAAATCAAGGTGTCGACTTCAAGGTCTTGTTGACATC
25 ACGAGACTCACAAGTTTGCATATGATGGGGGTTGAAGCTAATTCAATTA
TTAACGTGGGCCTTCTAACTGAAGCAGAAGCTCAAAGTCTGTTCCAACAA
TTTGTAGAACTTCTGAGCCCGAGCTCCAGAAGATAGGAGAGGATATCGT
AAGGAAGTGTTGCGGTCTACCTATTGCCATAAAAACCATGGCATGTACTC
TTAGAAATAAAAAGAAAGGATGCATGGAAGGATGCACTTTCGCGCATAGAG
30 CACTATGACATTCACAATGTTGCGCCCAAAGTCTTTGAAACGAGCTACCA
CAATCTCCAAGAAGAGGAGACTAAATCCACTTTTTTAATGTGTGGTTTGT
TTCCCGAAGACTTCGATATTCCTACTGAGGAGTTGATGAGGTATGGATGG
GGCTTGAAGCTATTTGATAGAGTTTATACGATTAGAGAAGCAAGAACCAG
GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTTGTTAATTGAAA
35 GTGATGATGTTGGGTGTGTCAAGATGCATGATCTGGTCCGTGCTTTTGT
TTGGGTATGTTTTCTGAAGTCGAGCATGCTTCTATTGTCAACCATGGTAA
TATGCCTGGGTGGCCTGATGAAAATGATATGATCGTGCCTCTTGCAAAA
GAATTTTCAATTAACATGCAAGGGTATGATTGAGATTCCAGTAGACCTCAAG
TTTCCTAAACTAACGATTTTGAACCTTATGCATGGAGATAAGTCGCTAAG
40 GTTTCCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCATGTTATATCAT
ACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCGATGCTCCACC
AACATTCGGGTGCTTCATCTCACTGAATGTTCAATTAAGATGTTTGATTG
CTCTTCTATCGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCAAATT

CTCACATTGAATGGTTACCTTCCACAGTCAGAAATTTAAAGAAGCTAAGG
T TACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAACAGGGTGTCTT
GAAAAGTTTTGTCAAACCTTGAAGAATTTTATATTGGAGATGCATCTGGGT
TTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTTACAACCTTTCT
5 GCATTAGAATTCGCGTTCTTTAATAACAAGGCTGAAGTGAAAAATATGTC
ATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGATGCTCTTTTGATG
AAAATATCAATATGAGTAGCCACTCATACGAAAACATGTTGCAATTGGTG
ACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGGTTATTTTTGAA
AACAGAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGATCTTGAAGATG
10 TTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCATTCTGCAATTTA
AAAGTTCTTATTATTTCAAAGTGTGTAGAGTTGAGATACCTTTTCAAAC
CAATCTTGCAAACACTTTGTCAAGACTTGAGCATCTAGAAGTTTGTGAAT
GTGAGAATATGGAAGAACTCATACATACTGGAATTGGGGGTTGTGGAGAA
GAGACAATTACTTTCCCTAAGCTGAAGTTTTTATCTTTGAGTCAACTACC
15 GAAGTTATCAAGTTTGTGCCATAATGTCAACATAATTGGGCTACCACATC
TCGTAGACTTGATACTTAAGGGCATTCCAGGTTTCACAGTCATTTATCCG
CAGAACAAAGTTGCGAACATCTAGTTTGTGTAAGGAAGGGGTAGATATATG
TTCTTTATGTTAATAACAATTTAAATAATATTTTCAACCAAATTTTCATAA
TATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGGCTATT
20 AAGGGATGATTATTTTGCAGGTTGTGATTCTTAAGTTGGAGACACTTCAA
ATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTTAGTGG
AGGTGAGAAAGTTAAGTTGAGAGCGATTAAAGTGAGTAGCTGTGATAAGC
TTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATCTTGAA
GAGCTTACAGTCGAGAATTGCGGTTCCATTGAGTCGTTATTCAACATTGA
25 CTTGGATTGTGTCGGTGCAATTGGAGAAGAAGACAACAAGAGCCTCTTAA
GAAGCATCAACGTGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGAGGATA
AAAGGTGCAGATAACTCTCATCTCATCAACGGTTTTCAAGCTGTTGAAAG
CATAAAGATTGAAAAATGTAAGAGGTTTAGAAATATATTCACACCTATCA
CCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAGGTTGC
30 GGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTTCAATTTCACTTTC
TTAATTAATTAAGGACTAAGCTCCTGTTTTTTGAATAATAAAGAGGTGGG
ATGACTAACTTGGGCATCACAATTGCAACAAAATGTTACAAACCATGAA
ACGTTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTTAAAAA
TATGGCTTAAATTTTTTTTATATTTATGTATCAACATGATTTTTTCATTAG
35 AGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATCAGAACTAATT
CTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTGTAGTG
AAATAAGCAACGGATTTAATAAGTTAACAACCTTAAATGTCATTTCTTAAC
AAAAAAACTATTTGGTTCAGAAAAACCGTAATTCAAGATAACTAAAATA
AAAATATTTGACATTCATAAGAGCATTTTTTTTTTCTAAATATGATTGCA
40 AATGAATAAACTTAAATTTATACAGAAAATTCTTTTATATATGTTATAC
AAAATTTACAAATTGAAATTGGATATGTTAATTAACGGTTTATAATTCTG
GTATCACAAGGGATATATAATAAATATTATTTTCTGTAGTCATTTGTA
ATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAGTTAA

ACTTTCATAAIAAAAATTTATAAATTATTTATTTTAAATAAATTAATTA
ATTAAGAGATATATCAAAAATTTAAAGTTATTATAACTTCAAATTTAACAA
TATAATTAGAAAAATATATGATCATAACTTTCTGCAACTCTTCTTTTGAT
TAAAATGACCAGAGAAGCTCTTAGTATATTTCTAATCAAAGTCTCAAAMC
5 TAATGAAGCATATAATTTGTGAAAATCAATTAGCATTAGGTTTTAAGAGT
CACCAAATTCAAAGAATAATCCAATGCTTTCATTACCACTATGGAGAAAA
TATTTTCTTAGTTTTAAATGAAATGAAACAACATTCAAACCTAATTGTTG
CTTATTAACCAAAGACCCATTACTTAGCCAAGAGTTTAACAAAAAAAAA
TTACATTCATGTATCATTATTCATGACTAGATATATATGAACATGAAGGG
10 AGTTTTTATAGAAAATATAATCATAGATATTCAACATAA CTTCAGGGAAT
TCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCAACCAAAG
AGAAGTTTAGCCTAGCATGGCTAAACTCAAGAACTAAAATAAGGATTAG
AAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTTGTTCTTG
ATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAATGCAAAGG
15 AGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAGCTGCACCAACCT
CCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCTGAAGGACC
CTATGTGGGTGCCTTGCGCGGGTGGAGCTGCATACGAAAGGTCTTTGGTC
TTTGTGAGGGTGATGTTGTGCGGGATAGCTTGTCGCATGCTCCGCGCGG
TTCACGCACATGTGCACAGGTGATGCATGGTGTGTGCGTTCTTGAGTTTT
20 GAGCCTCCGATGCTTAGTCCACTTG GCCCAATTTCGAGTCCAATCAGCTTA
TAACCCATTTTTCTTCAAGTTATCTTCAAGTTAAGCCCAATTG CCTTCT
CCAATCATCCATAA CTTCACAGAATCGCCCGTTCATCTTAATCCC GGAT
GCACAATTATTCTCCCGTCTTCATTTTAAGCAAGATACCACCTTCTTCAT
GCTTCATCCATCAATAGTACACTTCATGTATCATCTCTACTAGTTATTTA
25 GTCCACAATCCTTGTTGTCCTCCAAATTTAATTATCTCATT TAGTTCCCCG
TTCCGCTAGTTTTCTTAAAATTTGCAATTAAGCTCAGAGAAATATTAAGT
ACCCGAAATGGTCATAAAATAACAAAAAGGAAAATATGCATGAAGATTAA
CTAAATGATGAACGAAATATGCTAAAAATAGACTATAAAATGAAGTAAATA
AAATGAAATTATCGCACTCCGACCACCCTTATAGGCTTG TAGTCCACCCA
30 CCCTTCATTCTTG TACCAATATGGGATGGAAACATCATT AATTAAGCCA
AAAAGCTAACATATAAGGGTTTAGTGACAAAGGTAAGT ACTAAAGATGAA
AAT.AATCCATTTTTCTTGATATACACAACACACACATAGGGG CAGACGT
AGG.ATTTCAAAGTACAGATTGTTGGTGGCACATAAGTGTTGCTGGTGACA
TTTTTTTTTTTTTTTTTACGTAGTGGCACAACAGTAGGAAAA AACGAAAAAT
35 TCG-AAATTTTTTACAATTTGTCTAAAAAAAACAGTGGTTGTTGG TGCCAC
TATGGACACCAAAGTTGAACTGCCCCCCACGCGCGCACACACACACACA
CATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAAGAAAGAGAGAG
GAGAGTTTGGGATGTGATACTTCTTTTAGGAAATGGAGTTATATCTTTG
ATATTGTATTTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATA
40 AGTTTTATTTATTTTGATATGAAAAAAAAAAGTCTTTTATACATTGGATTT
AACATAAAAATCCAACAATATTAATCAAAAAGACCAMACATGTGGACAMW
TATGTATATAAWTAATTCACAATAGTCTTTAGGAATAGNATTATATATAT
AATTAATTCTCAATGGTCTTAGGAATAGTAAGTTCTTATATTTCAAACCT

TNGCCACAATTCTTTGKTTACTTWGACACTTYCCTCTCTCTAATTATATA
TATATATATATATATATATATATATATATACACACACACACACACACTAG
ATGTGTGCCCGCGCAAAGCAGTGACGTNNNGGAGAANACTTTCTTAAGCA
TAAATAATTATTATATTTTTTTATTGGGTATTATATAATAAAAAATTACAA
5 CTTTTAAATAAAATATTTATGTTTATACTTTATATTTATATTGCTTGTAT
ACTATTAATATAATAAATTAATATTTATGTCTAATTTATGAAATGTAAAT
TAATTTAAATACATGAATTTAATATTTTTTAAATTTTCAGTTTGCTTCAA
ATTGAGTTTCTTAATTATTTTTTTTTAATTCANGTATTCAAACCTTTGGTA
AGTATTAAGAATTATTTATGCATAATTGATTTATACAAAAAATTTGTA
10 ACTTATACATCTTAAATTTCAAGATATAACTAACATGTTTTACAATATAT
AT
TAAAGCGCAAAGGTCATAGGAATAGAATATTTTCTATTATTCTACGTTTT
GCCACAAAAGTTTGAACACTTTGCCACTTTTTGTCCCTCCTTAACCTTTT
CAATGTTTTGCGACAAAAGTTCCAAAACCTTTGCCACTTTGATCATTCCCTC
15 AACTTTTACCAGCAATTAGTTTGTGGAGTTGGCAGTTTGTATCCCCCTAA
CTTCGATATTCTCTACTGCTAGCCAAAAAGGGTTCCAGAGTTTCACACTT
TTGGTCCCTGACAGTAACCAAATGTGAGATGTCAAATTTTTGCCACATTA
GTTTGTGGAGTTGTCCCTTTTGGTCCCCCACCATTTCGATATTCTANTATA
CGACCTTATTTTTTNTCAAATAACAACACGTATATTTAATTACCAATTATA
20 GAAATAGATATCAAATAAAGTATTTGTAACACTGTGTAAGAACGGTGCTA
CTATAGGTAAAAATAAACATTTCAAAGTACGATATCCTAATTGGAAAAAG
AGTTTTAAAAAAATAACGACTAGGGGCGAGTTTTTTTTTACAAGTTTGAT
CAAATCATATCAAATTTAAGGTGGAACGGTGACCACATTAACCAGAAAT
GTAATTTATTCTTTGATTTTGATAATTTTTAATATTTTGTGTGATCTAT
25 GTATTTAAAGTAAACAACAAAGAACATAATCCAAAACCCTAAATTGCAA
GTCTCGCCCAATTTCTCTATCACTAGTCCTCACTTACGATGGCGTTACGT
CGCTCTCTCACTGCTTACAACCCTTTGTTGCTACTCATTACAATAACGAA
AAGTTGAATATCCATATATTTATTTGGATGTGGAATTGAACGAATCTCGT
CAAATTTTGTATTTTGTGATGGATTTGAGTAGAAGTTTGGGCAGAACGG
30 GAATGATGGTCTGCAAGTGGTTATAAACTTGATTCTGAGTTATTACTATA
TATGTAGCCTCTTTACAACGACCAAGGTTTCTTCCAGGTACCATTTGATC
TTTTTAGAACTTAGTTTTCTGAAACACCCTGATTTGGATCAAATATCACC
AACAACTCTTAAAACTTGATTAATCAATTGTTTTCTTCATCTTGATAAC
AAGTGGAATGATTTTCTACTTAGATTAACCTGAAAAAAAAGGTCCATGTG
35 CGTCTGGTGGATCTGGTAAATGAAGATGGAAGGGAGAGCTGACTTTAAAG
ACACAAACACGTCACCATATCTCTTATTTTATTTTAAATTTGCTTTTGGT
GTATTTTCTTTTTCTTATTTCTTTCTTTGATCTCCAGATGGTATGT
GGTGTGGATAATTTACACCTAGAGATTGGGAACGATGGGAAGGGGTCTGT
GATTTATGGCTGGCCGAGTTTTACTTATTAACCTCAATTTCAACCTAAATT
40 CTGATTCTTGTTTGAATAAAGTTGCATCTTTATTTTTGTATTATCTTGT
TGCATAGGATCCTTAGCATCTTTAATAATTTATTTGAAGGTGAAAGATC
CAACTATTTTTTAGCTGTTGGCATTTTCCATCATTGCAACTGTTTCTTG
AAAAAAAATACCTAAAATAAAAATAACCATTTTCAAATCCAAAATTATA

CAGGTATATATATATTTCTTTAATTGGCATCATCTAATTAAGAAAGATAT
CATTCCTGCCAAGTAAATTTACTTCAAACACATTCACACTGGTTTCAGTC
TAAGTTTATGTTGTTCTAAGAAGGCCAAAATGGGAAAGCAAGATAGGGAA
AAATAGTGTATTTTCAGTGGAAAGGGTATTTAGGCATTTTCTGTCAAAAG
5 TTGTTATTGCAGGCTTTTATGACCTGGAATCGTGTGTGGGAGGAGCATT
ATTATTCTGATTGCTTGTCTTTATCATTTTTTCTTAGCCTCTCGAAC
AGCTAGAAACCTTTTAATCTTTTGATTTTCAATGACGAAATTTTCCCT
GTTACTCCATTTGATTGTTGTTCTTCATGGTTCTAAGTGAGTTATTGGCT
CATCTGTTACTTCTTTTGATTGTTATTTTCATATCATGTTGTCCTTTGAA
10 TCAAGCTTTTCCATTTTCAACCAGGGCAAAAGGTCAAAAGTAACCTACTT
TATGAGATCAAAAACAGCAACCCATCGGATAACTTTTAGTTGGAGTTAAT
AGTTACAATTACCATTGTGATTAATAATTATAATATCTTGTATTAATTCA
TAAAAATTGGTACAGCACATATATGACATTTCAAAGGTTTTTGTGTTGACA
TATATATGCCTCTGGCGTTTTCTTTATTGGACTTGCAGACCTCATTCCAA
15 AGTTTATACGGTGACACCTTGGGCCCTGCTACTTCAGAAGGGACAACCTTG
GTCTTTTCATAACTTTATCGAATTAGATGTGGAAGGTAATCATGATGTTA
AAAAGATTATCCATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAG
ATT.AATGTAAGGTGGTGTAAAAGGGTAGAGGAGGTATTTGAAACTGCATT
GGAAGCAGCAGGGAGAAATGGAAATAGTGGAATTGGTTTTGATGAATCGT
20 CACAAACAACCTACCACTACTCTTGTCATCTTCCAAACCTTAGAGAAATG
AACTTATGGGGTCTAGATTGTCTGAGGTATATATGGAAGAGCAATCAGTG
GACAGCATTGAGTTTCCAAACCTAACAAGAGTTGATATCTATAAATGTA
AAAGGTTAGAACATGTATTTACTAGTTCCATGGTTGGTAGTCTATCGCAA
CTCCAAGAGCTACATATATCCAACCTGCAGTGAGATGGAGGAGGTGATTGT
25 TAAGGATGCAGATGATTCTGTAGAAGAAGACAAAGAGAAAGAATCTGATG
GGGAGACGAATAAGGAGATACTTGTGTTACCTCGTCTAAACTCCTTGATA
TTAAGAGAACTTCCATGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTT
TTC.ATTCCCATTATTGGATACTTTAAGAATTGAGGAATGCCCAGCAATAA
CCACCTTCACCAAGGGAAATTCCGCTACTCCACAGCTAAAAGAAATTGAA
30 ACACATTTTGGCTCGTTTTGTGCTGCAGGGGAAAAGACATCAACTCTCT
TATAAAGATCAAACAACAGGTAAATCAGATCTTTGTTGCTTTAATAATTC
TTAAACTACATTTGAAAAGCTTCATGCAAGTTTTTTTTGTTATATTGTCA
AAAACCGCAACCTACATTTTCAGCTTTATATTTATGTACTTTATGCAGGA
GTTCAAACAAGACTCTGATTAATGTGAAGTAAATACTAAAGGTAAATTAT
35 ATTTTCATGTTCCTAGTTGCCTATTAATTAATTGCCTTTTAGTTCATGAT
TTTTGGATGCATTCTTCATGATGATGTCAATCTTCTAATACCCCATTCAT
TGTTTGGTTGAATGTTGACTCTATGTCTTGATGAATATTCAAGGGAAGAA
TTGTTTCATCATATGAAGGACATTAAAGAAGAACATGGATGCTATGAAGAT
GTGGGAAAACAA

40

RG2B deduced polypeptide sequence (SEQ ID NO:90)

MSDPTGIAGAIINPIAQTALVPVTDHVGYMISCRKYVRVMQMKMTELNTSRISVEE
HISRNTRNHLQIPSQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLGGKAFKITE
QIESLTRQLSLISWTDDPVPLGRVGS MNASTSASLSDDFPSREKTFTQALKALEPNQK
5 FHMVALCGMGGVGKTRMMQRLKKAEEKKLFNYIVGAVIGEKTDPFAIQEAIADY
LGIQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGLSPFPNQ
GVD FKVLLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETSEPELQKIGE
DIVRKCCGLPIAIKTMAC TLRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQ
EEETKSTFLMCGLPEDFDIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQ
10 TNLLIESDDVGCVMHDLVRAFVLGMFSEVEHASIVNHGNMGPWPDENDMIVHSC
KRISLTCKGMIEIPVDLKF PKLTILKLMHGDKSLRFPQDFYEGMEKLVHVISYDKMKY
PLLPLAPRCSTNIRVLHLTECSLKMFD CSSIGNLSNLEVLSFANSHIEWLPSTVRNLK
KLRLDLRFCDGLRIEQGV LKSFVKLEEFYIGDASGFIDDNCNEMAERSYNLSALEF
AFFNKA EVKNMSFENLERFKISVGC SFDENINMSSH SYENMLQLVTNKGDVLD SK
15 LNGLFLKTEVLF LSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLISKCVELRYLFKL
NLANTLSRLEHLEVCECENMEELIHTGIGGCGEETITFPKLKFLSLSQLPKLSSLCHN
VNIIGLPHLVDLILKGIPGFTVIYPQNKLR TSSLLKEGVVIPKLET LQIDDMENLEEIW
PCELSGGEKVKLRAIKVSSCDKLVNLFPRNPMSLLHHLEELTVENCGSIESLFNIDLD
CVGAIGEEDNKSLLRSINVENLGKLREVWRIGADNSHLINGFQAVESIKIEKCKRFR
20 NIFTPITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEATGSISNLVFPSCLMH
SFHNLRLVLTLDNYEGVEVVFEIESESPTCRELVTTRNNQQPHLPYLQDLYLRNMD
NTSHVWKCSNWNKFFTLPKQQSESPFHNLT TINILKCKSIKYLFSPLMAELLSNLKDI
RISECDGIKEVVSNRDEDEEMTTFTSTHTTTTLFPSLDSLTL SFLENLKCIGGGGAK
DEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIEIVG CYALSSVIPCYAA
25 GQMOKL

RG2C polynucleotide sequence (SEQ ID NO:91)

ATAATATTACACAAAGGTAACGTCATTAATTAATTACGATACGAGACAGA
CTTTTTCAC TCGGACATNAACGGTCTATTCCTAACTTNANNTAATTNAAT
30 GAATTTAGGATGTGCTAATATGCATGTAANATTCGCTACCGTCATCTTTC
AAATGACCATATTTTATGTATTTATAATGAATCAATGAAAAACCGGATT
TCTATTTAAAATTCTTAAACTTCATCTTTTAAGCCAGGGTGAATACAAT
TGCTAGATCCACTGTTAATTTCCATCGAATTATGCCTGATCAATTGTTGG
CTGCCTACGATGCAGGTGCTACCACAAGAATATGGCCATGGAAACTGCTA
35 ATGAAATTATAAAACAAGTTGTTCCAGTTCTCATGGTTCCTATTAACGAT
TACCTACGCTACCTCGTTTCCTGCAGAAAGTACATCAGTGACATGGATTT
GAAAATGAAGGAATTA AAAAGAAGCAAAAGACAATGTTGAAGAGCACAAAGA
ATCATAACATTAGTAATCGTCTTGAGGTTCCAGCAGCTCAAGTCCAGAGC
TGGTTGGAAGATGTAGAAAAGATCAATGCAAAAGTGGA AACTGTTCTTAA
40 AGATGTCGGCTGTTGCTTCAATCTAAAGATTAGGTACAGGGCCGGAAGGG
ATGCCTTCAATATAATTGAGGAGATCGACAGTGT CATGAGACGACACTCT
CTGATCACTTGGACCGATCATCCCATTCTTTGGGAAGAGTTGATTCCGT

GATGGCATCCACCTCTACGCTTCAACTGAACACAATGACTTCCAGTCAA
GAGAGGTAAGGTTTAGTGAAGCACTCAAAGCACTTGAGGCCAACCACATG
ATAGCCTTATGTGGAATGGGGGGAGTGGGGAAGACCCACATGATGCAAAG
GCTGAAGAAGGTTGCCAAAGAAAAGAGGAAGTTTGGTTATATCATCGAGG
5 CGGTTATAGGGGAAATATCGGACCCCATTGCTATTCAGCAAGTTGTAGCA
GATTACCTATGCATAGAACTGAAAGAAAGCGATAAGAAAACAAGAGCTGA
GAAGCTTCGTCAAGGGTTCAAGGCCAAATCAGATGGAGGTAACACTAAGT
TCCTCATAATATTGGATGATGTCTGGCAGTCCGTTGATCTAGAAGATATT
GGTTTAAGCCCTTCTCCCAATCAAGGTGTGCGACTTCAAGGTCTTGTTGAC
10 TTCACGAGACGAACATGTTTGCTCAGTGATGGGGGTTGAAGCTAATTCAA
TTATTAACGTGGGACTTCTAATTGAAGCAGAAGCACAAAGATTGTTCCAG
CAATTTGTAGAACTTCTGAGCCCGAGCTCCACAAGATAGGAGAAGATAT
TGTTAGGAGGTGTTGCGGTCTACCCATTGCCATCAAAACCATGGCGTGTA
CTCTAAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCTCGTTTA
15 CAACACCATGACATTGGTAATGTTGCTACTGCAGTTTTTTAGAACCAGCTA
TGAGAATCTCCCGGACAAGGAGACAAAATCTGTTTTTTTGATGTGTGGTT
TGTTTCCCGAAGACTTCAATATTCTACCGAGGAGTTGATGAGGTATGGA
TGGGGCTTAAAGTTATTTGATAGAGTTTATACAATTATAGAAGCAAGAAA
CAGGCTCAACACCTGCATTGACCGACTGGTGCAGACAAATTTACTAATTG
20 GAAGTGATAATGGTGTACATGTCAAGATGCATGATCTGGTCCGTGCTTTT
GTTTTGGGTATGTATTCTGAAGTCGAGCAAGCTTCAATTGTCAACCATGG
TAATATGCCTGGGTGGCCTGATGAAAATGATATGATCGTGCACTCTTGCA
AAAGAATTTTCATTAACATGCAAGGGTATGATTGAGTTTCCAGTAGACCTC
AAGTTTCCTAAACTAACGATTTTGAACTTATGCATGGAGATAAAATCGCT
25 AAAGTTTCCTCAAGAATTTTATGAAGGAATGGAAAAGCTCCGGGTATAT
CATACCATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCAATGCTCC
ACCAACATTCGGGTGCTTCATCTCACGGAATGTTTCATTAAGATGTTTGA
TTGCTCGTGTATTGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCTA
ATTCTTGCAATTGAGTGGTTACCTTCCACGGTCAGAAATTTAAAAAAGCTA
30 AGGTTACTTGATTTGAGATTGTGTTATGGTCTCCGTATAGAACAGGGTGT
CTTGAAAAGTTTGGTCAAACCTTGAAGAATTTTATATTGGAAATGCATATG
GGTTTATAGATGATAACTGCAAGGAGATGGCAGAGCGTTCTTACAACCTT
TCTGCATTAGAATTCGCGTTCTTTAATAACAAGGCTGAAGTGAAAAATAT
GTCATTTGAGAATCTTGAACGATTTAAGATCTCAGTGGGATGCTCTTTTG
35 ATGGAAATATCAATATGAGTAGCCACTCATACGAAAACATGTTGCGATTG
GTGACCAACAAAGGTGATGTATTAGACTCTAACTTAATGGGTTATTTTT
GAAAACAGAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGATCTTGAAG
ATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCATTCTGCAAT
TTAAAAGTCCTTATTATTTCAAAGTGTGTAGAGTTGAGATACCTTTTCAA
40 ACTCAATGTTGCAAACACTTTGTCAAGACTTGAGCATCTAGAAGTTTGTA
AATGCAAGAATATGGAAGAACTCATACATACTGGGATTGGGGGTTGTGGA
GAAGAGACAATTACTTTCCCAAGCTGAAGTTTTTATCTTTGAGTCAACT
ACCGAAGTTATCAGGTTTGTGCCATAATGTCAACATAATTGGGCTACCAC

ATCTCGTAGACTTGAACTTAAGGGCATTCCAGGTTTCACAGTCATTTAT
CCGCAGAACAAGTTGCGAACATCTAGTTTGTTGAAGGAAGAGGTAGATAT
ATGTTCTTTATGTTAATACAATTTAAACAATATTTTCAACCAAATTTTCA
TAATATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGGCT
5 ATTAAGGGATGATAATTTTGAGGTTGTGATTCCCTAAGTTGGAGACACTT
CAAATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTTAG
TGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGATA
AGCTTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATCTT
GAAGAGCTTACAGTCGAGAATTGCGGTTCCATTGAGTCGTTATTCAACAT
10 TGACTTGGATTGTGTGCGGTGCAATTGGAGAAGAAGACAACAAGAGCCTCT
TAAGAAGCATCAACGTGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGAGG
ATAAAAGGTGCAGATAACTCTCATCTCATCAATGGTTTTCAAGCTGTTGA
AAGCATAAAGATTGAAAAATGTAAGAGGTTTAGAAATATATTCACACCTA
TCACCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAGGT
15 TGCGGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTCAATTTCACT
TTCTTAATTAATTANGGACTAAGCTCCTGTTTTTTGAATAATAAAGAGGT
GGGATGACTAACTTGGGCATCACAATTGCAACAAAATGTTACAAACCAT
GAAACGCTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTTAA
AAATATGGCTTAAATTTTTTTATATTTATGTATCAACATGATTTTTTCATT
20 AGAGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATTAGAACTAA
TTCTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTTAG
TGAAATAAGCAACGGATTTAATAAGTTAACAACCTAAATGTCATTTTCCTA
ACAAAAAACTATTTGGTTCAGAAAACTGTAATTCAAGATAACTAAAA
TAAAAATATTTGACATTCCTAAGAGCATTTTTTTCTAAATATGATTGCA
25 AATGAATAAACTTAAATTTATACAGAAAAGATTTTATATATGTTATAC
AAAATTTACAAATTGAAATTGGATATGTTAATTAACGGTTTATAATTCTG
GTATCACAAAGGGATATATAATAAAATATTATTTTTCTGTAGTCATTTAT
AATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAGTTA
AACTTTCATAATAAAAATTTATAATTATTATTTATTTTAAATAAATTATT
30 AATTAAGAGATATATCAAAAATTTAAAGTTATTATAACTTCAAATTTAAC
ATATAATTAAAAATATATGATCATAACTTTCCGCAACTCTTCTTTTGTA
TTAAATGACCAGAGAAGCTCTTAGTATATTTTCTAAATCAAAGTCACAA
AACTAATGAAGCATATAATTTTGTGAAAATCAATTAGCATTAGGTTTTAA
GAGTCACCAAATTCAAAGAGTAATCCAATGCTTTCATTACCACTATGGAG
35 AAAATATTTTCTTAGTTTAAATGAAATGAAAACAAACATTCAAACCTAATT
GTTGCTTATTAAACCAAGACCCATTACTTAGCCAAGAGTTTAACCAAAA
AAAATTACATTCATGTATCATTATTAATGACTAGATATATATGAATATGA
AGGGAGTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTCATG
GAATTCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCAACC
40 AAAGAGAAGTTTAGCCTAGCATGGCTAAACTCAAGAAAATAAAATAAGGA
TTAGAAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTTGTT
CTTGATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAATGCA
AAGGAGAGCCATTAAATTTCGTATGTATTGATCCCTTCAAAGCTGCACCA

ACCTCCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCTGAAG
GACCTATGCGGGTGCCTTGCGCGGGTGGAGCTGAATATGAAAGGTCTTT
GGTCTTTGTGAGGGTGTGTTGTGCGGGTTAGCTTGTGCGCATGCTTCCGC
GCGGTTGCGGCACATGTGCACAGGTGATGCATGGTGTGTACGTTCTTGAC
5 TTTTGAGCCTCCGATGCTTAGTCCACTTGGCCCAATTCGAGTCCAATCAA
CTTATGACCCATTTTTCTTCAAGTTATCTTCAAGTTAAGCCCAATTTGCC
TTCTCCAAATCATCCATAACTTCACAGAATCGCCCGTTCATCTTAATCCC
GAATGAACAATTATTCTCCCGTCTTCATTTTAAGCAAGATACCACCTTCT
TCATGCTTCATCCATCAATAGTACACTTCATGTATCATCTCTACTAGTTA
10 TTTAGTCCACAGTCCCTTGTTGTCCTCCAAATTTAATTATCTCATTAGTT
CCCGTTCCGCTAGTTTCCTTAAAATTTGCAATTAAGCTCACAGAAATATT
AAGTACCCGAAATGGTCATAAAATAACAGAAAGGAAAATATGCATGAAGA
TTAACTAAATGATGAACGAAATATGCTAAAATAGACTATAAAATGAAGTA
AATAAAATGAAATTATCGCACTCCGACCACCCTTATAGGCTTGTAGTCCA
15 CCCACCCTTCATTCCCTTGTAACCAATATGGGATGGAAACATCATTAAATA
GCCAAAAAATAACATATAAGGGGTGAGTGACAAAGGTAAGTACTAAAGA
TGAAAAAATCCATTTTTCTTGTATATACACAACACACACATAGGGGCAG
ACGTAGGATTTCATAGTACAGATTGTTGGTGGCACATAAGTGTTGCTAGT
GACATTTTTTTTTCTTTTACGTAGTGGCACAACAGTARAAAAAACRAAA
20 AATTTCGAAATTTTTTACAATGTGCCTAAAAAAACAGTGGTTGTTGGTGC
CACTATGGACACCAAAGTTGAACTGCCCTGCGCGCGCACACACACAC
ACATAAAGTTTG
GGATGTGATACTTCTTTTGGGAAAATGGAGTTATATCTTTGATATTGTAT
TTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATAAGTTTTATT
25 TATTTKGATATGAAAAAAAAGTCTTTTATACATTGGATTTAACATAAAA
ATCCAACAATATTAATCAAAAAGACCAAACATGTGGACAATTATGTATAT
AATTAATTCACAATAGTCTTTAGGAATAGNATTATATATATAATTAATTC
TCAATGGTCTTAGGAATAGTAAGTTCTTATATTTCAAACNTTTGCCACAN
TTCTTTGNTTACTTNGACACTTTYCTCTMWNNANWMWWTWATATATATAT
30 ATATATATATATAHAHAHAHAVACACACACACTAGATGTGTGCCMGCGCA
AAGCAGTGACGTNNNGAGAANACTTTCTTAAGCATAAATAATTATTATA
TTTTTTATTGGGTATTATATAATAAAAAATTACAACTTTTAAATAAAATA
TTTATGTTTATACTTTATATTTATATTGCTTGTATACTATTAATATAATA
AATTAATATTTATGTCTAATTTATGAAATGTAAATTAATTTAAATACATG
35 AATTTAATATTTTTAAAATTTTCAGTTTGCTTCAAATTGAGTTTCTTAAT
TATTGACCAAACATGTGGACAATTATGTATATAATTAATTCACAATAGTC
TTTAGGAATAGTATTATATATATAATTAATTCTCAATGGTCTTAGGAATA
GTAAGTTCTTATATTTCAAACTTTTGCCACAATTCTTTGCTTACTTTGAC
ACTTTTCCTTCCTAACTTTACATATATATATATATATTAAGCGCAAAGGTC
40 ATAGGAATATAATATTTTCTATTATTCTACGTTTTGCCACAAAAGTTGA
ACACTTTGCCACTTTTTGTCCCTCCTTAACCTTTTCAATGTTTTGCGACA
AAAGTTCCAAAACCTTTGCCACTTTGATCATTCCTCAACTTTTCACCGCAT
TAGTTTGTGGAGTTGGCAGTTTTGGTCCCTCTAACTTCGATATTCTCTAC

TGCTAGCCAAAAAGGGTTCCAGAGTTTCACACTTTTGGTCCCTGACAGTA
ACCAAATGTGAGATGTCAAATTTTTGCCACATTAGTTTGTGGAGTTGTCC
CTTTTGGTCCCCCACATTCGATATTCTACTATACGATCTTATTTTCTC
AAATAACAACACGTATATTTAATTACTAATGATAGAAATAGATATCAAAT
5 AAAGTATTTGTAACACTGTGTAGAGTTTTTTTTTACAAGTTTGTATCAAA
TCATATCAAAATTTAAGGTGGAACGGTGACCACATTAACCAGAAATGTAA
TTTATTCTTTGATTTTGATAATTTTAATATTTTGTGTGATCTATGTAT
TTAAAAGTAAACAACAAGAACATAATCCAAAACCCTAAATTGCAAGTCT
CGCCCAATTTCTCTATCACTAGTCCTCACTTACGATGGCGTTACGTCGCT
10 CTCTCACTGCTTACAACCCTTTGTTGCTACTCATTACAATAACGAAAAGT
TGAATATCCATATATTTATTTGGATGTGGAATTGAACGAATCTCGTCAAA
TTTTTGATTTAGTTGATGGATTTGAGTAGAAGTTTGGGCAGAACGGGAAT
GATGGTCTGCAAGTGGTTATAAACTTGATTCTGAGTTATTACTATATATG
TAGCCTCTTTACAACGACCAAGGTTTCTCCAGGTACCATTTGATCTTTT
15 TAGAACTTAGTTTTCTGAAACACCCTGATTTGGATCAAATATCACCAACA
ACTCTTAAAAACTTGATTAATCAATTGTTTACTTCATCTTGATAACAAGT
GGAATGATTTTCTACTTGAAAAAAAAGGTCCATGTGCGTCTGGTGGATCT
GGTAAATGAAGATGGAAGGGAGAGCTGACTTTAAAGACACAAACACGTCA
CCATATCTTTTATTTTATTTTAAATTTTCTTTTTTCTTATTTCTTTCTTT
20 CTTGATCTCCAGATGGTATGTGGTGTGGATAATTTACACATAGAGATTGG
GAACGACTGTGATTTAGAGAGGACGTGGCTTGGGGTTGAGGATGGTTTAT
GGCTGGCCGAGTTTCATTTATATAAAACAAACAATATATAAAACAAGGGG
TAAATGGCCATCTTATATGTATTTAACCGTCCTCTTTTTTTTTTTTTTT
TTTTTTTTTTTTTTTTTTTGTAAATTAAGAAGGGGTATACCAGTGTCAAGC
25 CTCTTATTCCTAACCAAGTCAAATAGGGACTTAGGTTGTTTGGAAACAGTT
CCGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAACTTAACCCTT
CAATTAACCTACCTTTTTCTTATTAACCTCAATTTCAACCTAAATTCTGAT
TCTTGTTTGAATAAAGTTGCATCTTTATTTTGTATTATCTTGTTGCAT
AGGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAAAGATCCAAC
30 ATTTTAAATCTGTTGACGTTTCCATCATTTGCAACTGTTTCTTGAAAAA
AAAATACCTAAAATCAAATAACCATTTTCAAATCCAAAATTATAAGAGA
GAATTGTAAATGGACATGGAATCATAAATCATTAAACACAGTTCAGTAAAC
AAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAA
AGAGACATTACAAGAAGCCACTGGCAGTATTCAAATCTTGATTCCCAT
35 CCTGTCTCATGCACTCTTTTCATAACCTCCGTGTGCTTACATTGGATAAT
TATGAAGGAGTGGAGGTGGTGTGTTGAGATAGAGAGTGAGAGTCCAACAAG
TAGAGAATTGGTAACAACCTCACAATAACCAACAACAGCCTATTATACTTC
CCTACCTCCAGGAATTGTATCTAAGGAATATGGACAACACGAGTCATGTG
TGGAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATC
40 AGAATCACCATTCCACAACCTCACAACCATAGAAATGAGATGGTGTGTCATG
GCTTTAGGTACTTGTTTTCGCCTCTCATGGCAGAACTTCTTTCCAACCTA
AAGAAAGTCAAGATACTTGGGTGTGATGGTATTAAAGAAGTTGTTTCAAA
CAGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACAAAA

CCACCAACTTGTTCCCTCATCTTGATTCTCTCACTCTAAACCAACTGAAG
AATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGA
AATATCTTTCAATAATACCACTGCAACGACTGCTGTTCTTGATCAATTTG
AGGTATGCTTTGTACATATTCAATTATTTATTTAATTTTCCTTTTTTATTT
5 GCAATATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAA
TTACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAACTC
TAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGA
TGGGTAATATAGGCAATTTAAGTTTTATTTCTGTTAAAGCAGTATTTAGC
AAGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGT
10 CATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATTAGGGGTTA
TCAGGTGACAGATATTGTAGAATAGAACAATATGTAATATTACCCAAAAC
TATTTTTTCTAAGGTTGCTCTGTTAAATATGTGCTTTCTTGATTTCATTG
AATTTGCATTCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAA
TCCCGAAATTAATTAACAAAAAACAACAAAAGTAAATTTTTGATATGGA
15 GAGCACTGGTATCATTTAGTATATAAAAAAACTAGATTTTGAATTAAGTT
TCTTATATAAAAGCTGTGTATATAGTTTAATTAGTTTTACATCATTTTTTC
CATGTGGTGTGTCAGTTGTCTGAAGCAGGTGGTGTCTTGGAGCTTATG
CCAATACGCTAGAGAGATAAAAAATAGGCAACTGCCATGCATTGTCAAGTG
TGATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGAGA
20 GTAATGGCTTGCAATGGGATGAAGGAGGTATTTGAAACTCAATTAGGGAC
GAGCAGCAACAAAAACAACGAGAAGAGTGGTTGTGAGGAAGGAATTCCAA
GAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTAAGTATT
GGAATTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGAAAG
CCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTGCTACAGAATGAAAG
25 TGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACAACA
ACAACAACGAAGGGGGCATCTTCTTCTTCTTCTTCTTCTAAGAAGGTGGT
GGTCTTTCCTTGTCTAAAGTCCATTGTATTGGTCAATCTACCAGAGCTGG
TAGGATTCTTCTTGGGGATGAATGAGTTCGGGTTGCCTTCATTAGATAAA
CTTATCATCGAGAAATGCCCAAAATGATGGTGTTTACAGCTGGTGGGTC
30 CACAGCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACTC
TTGATCAAGAATCTGGCCTTAACCTTTCATCAGGTACATATATATTCCTTT
AATTGGCATCATCTAATTAAGAAAGATATCATTCCTGCCAAGTAAATTA
CTTCAAACACATTCACACTAGTTTTCAGTCCAAGTTTATGTTGTTCTAGGA
AGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGAGTATTTTCAGTGGA
35 AGGGTATTTTAGGTATTTTCTGTCAAAAATTGTTATTGCAGGCTTTTTAG
TACCTGGAAGAGCATGATTATTCTCGATTGCTTGTCTTTATCATTTT
TCTTAGCCTAGCATGATTTTCAATGAAATCTTCCCTGTTACTCCATTG
ATTGTTGTTCTTCATGGTTCTAAGTGAGTTAGTGGCTCATCTGTTACTTC
TTTTGATTGTTATTTTCATAGCATGTTGTCACTTGAATCAAGCTTTTCCA
40 TTTTCAACAAGGACAAAAGGTCAAAACTAACCTACTTTATGAGATCAAAA
ATAGCAACCCATCGGATAACTTTTAGTTGGAGTTAATACTTACAATTACC
ATTGTGATTAATAATTATAATATCTTGTATTAATTCATAAAAATTGGTAC
AGC.ACATATATGACATTTCAAAGGTTTTTGTGTTGACATATATATGCCTCT

GGCGTTTTCTTTATTGGACATGCAGACTTCATTCCAAAGTTTATACGGTG
ACACCTTGGGCCCTGCTACTTCAGAAGGGACAACCTGGTCTTTTCATAAC
TTTATCGAATTAGATGTGAAATCTAATCATGATGTTAAAAAGATTATTC
ATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGTAAAGATTAATGTAATGT
5 GGTGTAAGGGTAGAGGAGGTATTTGAACTGCATTGGAAGCAGCAGGG
AGAAATGGAAATAGTGAATTGGTTTTGATGAATCGTCACAAACAACCTAC
CACTACTCTTGTCAATCTTCCAAACCTTGGAGAAATGAAGTTACGGGGTC
TCGATTGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTGAG
TTTCCAAACCTAACAAGAGTTGAAATTTATGAATGTAATTCATTAGAACA
10 TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAG
AGATTGGTTTTGTGCAACCATATGGAGGTCGTGCATGTTTCAGGATGCAGAT
GTTTCTGTAGAAGAAGACAAAGAGAAAGAATCTGATGGCAAGATGAATAA
GGAGATACTTGTGTTACCTCATCTAAAGTCATTGAAATTACTACTTCTTC
AAAGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATT
15 TTGGATACTTTGGAATCTACGAATGCCAGCAATAACCACCTTCACCAA
GGGAAATTCCGCTACTCCACAGCTAAAAGAAATGGAAACAAATTTTGGCT
TCTTTTATGCTGCAGGGGAAAAAGACATCAACTCCTCTATTATAAGATC
AAACAACAGGTAAACCAGATCTTTGTTGCTTTAATAATTCTTAAACTACA
TTTGAAGAGCTTCATGCAAGTTTTTTTTGTTATATTGTCAAAAACCGCAA
20 CCTACATTTTCAGCTTTATTTATGTACTTTATGCAGGATTTCAAACAA
GACTCTGATTAATGTGAAGTGAATATTAAGGTAAATTATATTTTCATGT
TCCTAGTTGCCTATTAATTAAGGCCTTTTAGTTCTGTGATTTTTGGATGT
ATTCTTCATGATGATGTCAATCTTCTAATACCCCATTCATTGTTTGGTTG
AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTTCATCA
25 TATGAAGGACATTAAAGAACATGGTGCTAT

RG2C deduced polypeptide sequence (SEQ ID NO:92)

MAMETANEIHKQVVPVLMVPINDYLRYLVSCRKYISDMDLKMKEKAKDNVEEH
KNHNISNRLEVPAAQVQSWLEDVEKINAKVETVPKDVGCCFNLKIRYRAGRDAFNI
30 IEEIDSVMRRHSLITWTDHPIPLGRVDSVMASTSTLSTEHNDFQSREVRFSEALKALE
ANHMIALCGMGGVGKTHMMQRLKKVAKEKRKFGYIIEAVIGEISDPIAIQQVVADY
LCIELKESDKKTRAELRQGFKAASDGGNTKFLIILDDVWQSVLEDIGLSPSPNQG
VDFKVLTSRDEHVCSVMGVEANSINVGLLIEAEAQRLLFQQFVETSEPELHKIGEDI
VRRCCGLPIAKTMACTLRNKRKDAWKDALSRHQHDIGNVATAVFRTSYENLPD
35 KETKSVFLMCGLPEDFNIPTEELMRYGWGLKLFDRVYTHEARNRLNTCIDRLVQT
NLLIGSDNGVHVKMHDLVRAFLGMYSEVEQASIVNHGNMPGWPDENDMIVHSC
KRISLTCKGMIEFPVDLKFPLTILKLMHGDKSLKFPQEFYEGMEKLRVISYHKMKY
PLLPLAPQCSTNIRVLHLTECSLKMFDSCIGNLSNLEVLFSFANSIEWLPSTVRNLK
KLRLDLRLCYGLRIEQGVLSLVKLEEFYIGNAYGFIDNCKEMAERSYNLSALEF
40 AFFNKAIEVKNMSFENLERFKISVGCSPDGNINMSSHSYENMLRLVTNKGDVLDLSD
LNLGLFKTEVLFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLIIKSCVELRYLFLK
NVANTLSRLEHLEVCKCKNMEELIHTGIGGCGETITFPKLKFLSLSQLPKLSGLCH

NVNIIGLPHLVDLKLKGIPGFTVIYPQNKLRSSLLKEEVVIPKLETLOIDDMENLEEI.
WPCELSGGEKVKLREIKVSSCDKLVNLFPRNPMSLLHHLEELTVENC GSIESLFNID
LDCVGAIGEEDNKSLLRSINVENLGKLRVWRIKGADNSHLINGFQAVESIKIEKCK
RFRNIFTPITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEATGSISNLVFPSC
5 LMHSFHNLRVLTLDN YEGVEVVFEIESESPTSRELVTTHNNQQPIILPYLQELYLR
NMDNTSHVWKCSNWNKFFTLPKQQSESPFHNLTTIEMRWCHGFRYLFSPMAELL
SNLKKVKILGCDGIKEVVSNRDDEDEEMTTFTSTHKTTNLFPHLDSLTLNQLKNLK
CIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIKIGNCHAL
SSVIPCYAAGQMQLQVLRVMACNGMKEVFETQLGTSSNKNNEKSGCEEGIPRVN
10 NNVTMLPNLKILSIGNCGGLEHIFTSALESRLQLQELTIKGCYRMKVIVKKEEDEY
EQQTTTTTTKGASSSSSSSKKV VVFPCLKSIVLVNLP ELVGFFLGMNEFRLPSLDKLI
EKCPKMMVFTAGGSTAPQLKYIHTRLGKHTLDQESGLNFHQTSFQSLYGD TLGPAT
SEGTTWSFHNFIELDVKSNDHVKKIIPSELLQLQKLVKINVMWCKRVEEVFETALE
AAGRNGNSGIGFDESSQTTTTLVNLPNLGEMKLRGLDCLRYWKSQWTA FEFPN
15 LTRVEIYECNSLEHVFTSSMVGSLLQLQELEIGLCNHMEVVHVQDADVSVEEDKEK
ESDGKMNKEILVPLHLKSLKLLLLQSLKGFSLGKEDFSFPLD TLEIYECPAITTF TK
GNS.ATPQLKEMETNFGFFYAAGEKDINSSIIKIKQQDFKQDSD.

RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94)

20 ACGACCACTATAGGGCGAATTGGGCCCCGACGTCGCATGCTCCCGGCCGCC
ATGGCCGCGGGATGTAAAACGACGCCAGTCGAATCGTAACCGTTCGTAC
GAGAATCGCTGTCCTCTCCTTCAACCATTTAATGTATATGAGCTAAATTG
AAACATCTACTATCATGTTTAAATTTATAAACTTTTCTTTAGATTAC
TTGTCTGGATGTGTTAATAAAACCCAATTTCCCATATGCGTAGAGATCA
25 TAG.ATGTA ACTATTGTTAATCAATTTGCCTGCCAAGTTTAATAATTAT
ACTTGGATATTAACAAAACCTTTATCTAACGACCAAGGTAATATTA AAAAT
AGGTTATTATTCTTCATGCTAATTAAGATGGGTTGCAAAAGTGAGACC
ATG.AAAACATTAACACGTTGATATTTTCAACTTTTATTCTTT CATATTCA
CCATATTTTTTACTTTTCGTATTGATTAATCATCTTTCAATCACAGGCTCC
30 TTGGCAAAAAGTCAGATCTATTAACAAATACTTCCATGTGGTTGCAAATT
ACAAGGATTTCAACATAATTACCAAAACATAGCATTATCATAAGATCGAA
TAATAATCAAATTCTTCTATAATATTACACAAAGGTAACGTCATTAATTA
ATTACGATACGAGACAGACTTTTTCACTCGTGACATCAACGGTCTATTCT
AACTTTACTTAATTAATGAATCTAGGATGTGCTCATATGCATGTAATAT
35 TTGCTACCGTCATCTTTCAAATGACCATATTTTATGTATTTATAATGAA
TCAATGAAAAACCGGATTTCTATTTAAATTTCTTAAACTTCATCTTTTA
AGCCAGGGTGAATACAATTGTAGATCCACTGTTAATTTCCATCGATTATG
CGTGATCAATTGTTGGCTGCATACGATGCAGGTGCTACCACAAGAATATG
GCCATGGAACTGCTAATGAAATTATAAAACAAGTTGTTCCAGTTCTCAT
40 GGTTCCTATTAACGATTACCTACGCTACGTCGTTTCCTGCAGAAAGTACA
TCAGTGACATGGATTTGAAAATGAAGGAATTAAGAAGCAAAAGACAAT
GTTGAAGAGCACAGAATCATAACATTAGTAATCGTCTTGAGGTTCAGC

AGCTCAAGTCCAGAGCTGGTTGGAAGATGTAGAAAAGATCAATGCAAAAG
TGGAAGTGTTCCTAAAGATGTCGGCTGTTGCTTCAATCTAAAGATTAGG
TACAGGGCCGGAAGGGATGCCTTCAATATAATTGAGGAGATCGACAGTGT
CATGAGACGACACTCTCTGATCACTTGGACCGATCATCCCATTCTTTGG
5 GAAGAGTTGATTCCGTGATGGCATCCACCTCTACGCTTTCAACTGAACAC
AATGACTTCCAGTCAAGAGAGGTAAGGTTTAGTGAAGCACTCAAAGCACT
TGAGGCCAACCACATGATAGCATTATGTGGAATGGGGAGAGTGGGGAAGA
CCCACATGATGCAAAGGCTGAAGAAGGTTGCCAAAGAAAAGAGGAAGTTT
GGTTATATCATCGAGGCAGTTATAGGGGAAATATCGGACCCCATTTGCTAT
10 TCAGCAAGTTGTAGCAGATTACCTATGCATAGAGCTGAAAGAAAGCGATA
AGAAAACAAGAGCTGAGAAGCTTCGTCAAGGGTTCAAGGCCAAATCAGAT
GGAGGTAACACTAAGTTCCTCATAATATTGGATGATGTCTGGCAGTCCGT
TGATCTAGAAGATATTGGTTTAAAGCCCTTCTCCCAATCAAGGTGTCGACT
TCAAGGTCTTGTTGACTTCACGAGACGAACATGTTTGCTCAGTGATGGGG
15 GTTGAAGCTAATTCAATTATTAACGTGGGACTTCTAATTGAAGCAGAAGC
ACAAAGATTGTTCCAGCAATTTGTAGAACTTCTGAGCCCGAGCTCCACA
AGATAGGAGAAGATATTGTTAGGAGGTGTTGCGGTCTACCCATTGCCATC
AAAACCATGGCGTGACTCTAAGAAATAAAAGAAAGGATGCATGGAAGGA
TGCACCTTCTCGTTTACAACACCATGACATTGGTAATGTTGCTACTGCAG
20 TTTTGTAGAACCAGCTATGAGAATCTCCCGGACAAGGAGACAAAATCTGTT
TTTTTGATGTGTGGTTTGTTCCTCGAAGACTTCAATATTCCTACCGAGGA
GTTGATGAGGTATGGATGGGGCTTAAAGTTATTTGATAGAGTTTATACAA
TTATAGAAGCAAGAAACAGGCTCAACACCTGCATTGAGCGACTGGTGCAG
GCAAATTTACTAATTGGAAGTGATAATGGTGTACACGTCAAGATGCATGA
25 TCTGGTCCGTGCTTTTGTTCCTGAGGATGTATTCTGAAGTCGAGCAAGCTT
CAATTGTCAACCATGGTAATATGCCTGGGTGGCCTGATGAAAATGATATG
ATCGTGCCTCTTGCAAAAGAATTTCAATTAACATGCAAGGGTATGATTGA
GATTCCAGTAGACCTCAAGTTTCCTAAACTAACGATTTTGAACTTATGC
ATGGAGATAAGTCTCTAAAGTTTCCTCAAGAATTTTATGAAGGAATGGAA
30 AAGCTCCAGGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTT
GGCACCTCAATGCTCCACCAACATTGCGGTGCTTCATCTCACTGAATGTT
CATTAAAGATGTTTGATTGCTCTTCTATCGGAAATCTATCGAATCTGGAA
GTGCTGAGCTTTGCTAATTCTCGCATTGAATGGTTACCTTCCACAGTCAG
AAATTTAAAGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCC
35 GTATAGAACAGGGTGTCTTGAAAAGTTTGGTCAAACCTTGAAGAATTTTAT
ATTGGAAATGCATATGGGTTTATAGATGATAACTGCAAGGACATGGCAGA
GCGTTCTTACAACCTTTCTGCATTAGAATTCGCGTTCTTTAATAACAAGG
CTGAAGTGAAAAATATGTCATTTGAGAATCTTGAACGATTCAAGATCTCA
GTGGGGTGCTCTTTTGATGGAAATATCAGTATGAGTAGCCACTCATACGA
40 AAACATGTTGCAATTGGTGACCAACAAAGGTGATGTATTAGACTCTAAAC
TTAATGGGTTATTTTGAACAGAGGTGCTTTTTTAAAGTGTGCATGGC
ATGAATGATCTTGAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTC
CTCTTCATTCTGCAATTTAAAGTCCGTATTATTTCAAAGTGTGTAGAGT

TGAGATACCTTTTCAAACCTCCATGTTGCAAACACTTTGTCAAGCCTTGAG
CATCTAGAAGTTTGTGGATGCGAAAATATGGAAGAACTCATACATACTGG
GATTGGGGGTTGTGGAGAAGAGACAATTACTTTCCCAAGCTGAAGTCTT
TATCTTTGAGTCAACTACCGAAGTTATCAGGTTTGTGCCATAATGTCAAC
5 ATAATTGGGCTACCACATCTCGTAGACTTGAACTTAAGGGCATTCCAGG
TTTCACAGTCATTTATCCGCAGAACAAGTTGCGAACATCTAGTTTGTGTA
AGGAAGAGGTAGATATATGTTCTTTATGTTAATACAATTTAAATAATATT
TTCAACCAAAATTTTCATAATATATCTGTAATTTGATTGTATGATGTGTTA
TTGTTTATATGTGGCTATTAAGGGATGATTATTTTGCAGGTTGTGATTCC
10 TAAGTTGGAGACACTTCAAATTGATGGCATGGAGAACTTAGAAGAAATAT
GGCCTTGTGAGCTTAGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAA
GTGAGTAGCTGTGATAAGCTTGTGAATCTATTTCCGCACAATCCCATGTC
TCTGTTGCATCATCTTGAAGAGCTTAAAGTCAAAAATTGTCGTTCCATTG
AGTCGTTATTCAACATCGACTTGGATTGTGTCAGTGCAATTGGAGAAGAA
15 GACAACAAGAGCATCTTAAGAAGAATCAAAGTGAAGAATTTAGGGAAGCT
AAGAGAGGTGTGGAGGATAAAAAGGTGCAGATAACTCTCGTCCCCTCATCC
ATGGCTTTCCAGCTGTTGAAAGCATAAGTATCTGGGGATGTAAGCGGTTT
AGAAATATATTCACACCTATCACCGCCAATTTTGATCTGGTGGCACTTTT
GGAGATTACATAGGAAATTACAGAGAAAATCATGAATCGGAAGAGCAGG
20 TAACGCTTTCAATTTCACTTTCTTACTTAATTAAGGACTAAGCTCTTGTT
TTTTGAATAATAAAGAGGTGGGATGACTAACTTGGGCATCACAATTGTA
ACAAAATGTTACAAACCATGAACGTACAAACCATTTCTTGAATTAAGGTT
TCAATACAAGTCATTTACAAATATGGCTTAAGTTTTTTTATATTTATGTA
TCAACATTATTTTTTCATTAGAGGTCAATTATTATAATAGTAAGTTTAAAGC
25 AATTTAAATTAGCACTAATTTTTTCATCATCTAACTTTAGCTAATAAATCG
TTATAAATGTCAATAGCTAAAATAAAAATATTTGACATTCCTGAGAGCA
ATTTTTTCTAAACATGATTGCAAATGATTAAACTTAAATTTAAACTAAA
AAGATTTTTATATATGTTATACAAAATTTACAAATTGAAATTGGATATGT
TAATTAACAGTTTATAATTATTGTATTACAAAGCGATATATAATAAAATA
30 TTATTTTTCTGTAGTCATGTATAATTGTATATGTAAATGATTTTTTAAGA
TGGTAGAAGTGGAAACTAGTCAATCTCACTTAACTCATTGTACACCAGT
TTTATATCCGTTTCTCTCTCTCTCTCTTCTTGCCCTCCATCTTTTTTCAAC
TCATAACACATAAAAATAACATATTTTCCAACACATTTAAGTCACTACCA
CATCATTATTTTTAATTTAATTAATTAAGTAAATATAAAATTAATAAAAA
35 CATAACATTTTTTTTATTAAAAGGCACTAATACAAATAAAAAAGATACACGG
TAAATAAAAAAACGATAATTAGAAAAAAACATAATAAAAAAAGACAACA
TTAAAAATAWAAAGCGACAATAAATTAATAATGATCAAGAAAATTCT
AAAACCTCCACCATATTTTTCTGCAATTTGTCAATTTATGTTCAAACACCA
TTCGCAGAATCCCTCCTATCAAGTGATCATGTTGATTGAGAAAAAACTGT
40 ATGTCTCTCTCATGTATCTCCAAGTCCAACAAGTTAGCTTTCATTTCTTC
ATTTTCTCATGTAAGACGCAAATTTTCATCCCGATATTGTTTTCTATCTT
CCACCTCTACTTTATTACAGTGTGGATGAAGGAGAGGACAGCGATTCTC
GTACGAACGGTTACGATTGCGACTGGCCGTCGTTTTACAATCCCGCGGCCA

TGGCGGCCGGGAGCATGCGACGTCGGGCCCATTTCGCCCTATAGTGGTCGT
AATACA (SEQ ID NO:93)

Sequence gap

5 TGAGCCTCCGATGCTTAGTCCACTTGGCACAGTTCAAGTCCAATCAACTT
ATAACCCATTTTCTTCAAGTTGTCTTCAAGTTAAGCCCAATTTGCCTTC
TCCAAATCATCCATAACTTCATGGAATCGCCCCTTCATCTTAATCCCGAA
TGCACAATTATTCTCCCATCTTCATTTTAAGCAAGAGGCCACCTTCTTCA
TGCTTCATCCATCAATAGTCTGTTGGAATAGTGTCTAAGGCTGCAACTAT
10 ATTAGACAAGTATTTGACCCGTTGTGCATGGTCCTTTTGGGTTCCTTC
ACCATAGCAACTTGATAGGATGATTTATTAAGAGAGAGTAAATATTATTA
ATATATTATGAGAATAATATAATGAATAATATATTTGTTATTTGATTAAT
ATAAGTCATAGAATTAATTAGAATTAATTTGGTGACTTAAAGAGATTAAT
TAAATAAAGGGGTATAAACTGTCAATTGTTTGATAGTTAAGCTTTAGACT
GTAATCCATTTGGATATGGTATGGACGAATCCTAAGGGATTTAGGATAG
15 CTAAAATCGTCCATATGAGTTATCTAAGAAGGATTTGGATAGCCTTAAGA
GAAGATTATCTGATAGGGACTTATCTGTAATCCTTAAGGAGTCTACAAGT
ATAAATAGACCCTATGGCTGATGGAATTCGACACATCTCCTAAAGTAAGA
GAGCCTTGGCCGAATTCCTCCCCTCACCTCTCTCCTAAATCATTCTTCTT
GCTATTGGTGTTTGTAAAGCCATTAGAGGAGTGACATTTGTGACTCTAGAA
20 TCTCCAAGACCTCAAGATCAACAAGGAATTCAAAGGTATGATTCTAGATC
TGTTTCAATGTTGTTATTTGTCCTAATTAGTCATTAGAAGACTTGGATTC
AAAGCATGTTTATTAGAAAGCCTAGATCYGAGCAATAGGGTTTTGCATGC
GCACATAGGAAAGTTCTTATGGCTAAAACCCATCATAGTCCACTTCATGT
ATCATCTCTACTAGTTATTTAGTCCATAATCCTTGTTGTCCTCCAAGTTT
25 AATTACCTCCCTTAGTTCCTGTTCTGCTAGTTTCCTTAAAATTTGCTATT
AAGATCACAGAACTAGAGAGTACCCAAAATGGTTATAAAATAACAAAAAG
GAAAATATGCATGAAGATTAATAATTATAAATGTAATATGCTAAAATA
AACTATAAAAAAAAAGTAAATAAAATGAACTATCACACTCCGACCACCC
TTATAGGCTTGTACTGCACCCACCCTTCATTCCTTGTACCAATATGGGAT
30 GGAAACATTATTCATTAAGCCAAAAAACTAACATTTAAGGGGTGAGTGAC
AAAGGTAAGTACTAAAGACAACAATAATCCATTTTCTTGTACATACACA
ACACACACATAGGGGCGGACGTAGGATTTGTAGTATGTGTTGTGGGTGAC
ACATTTTTTCTTTTACGTAGTGACACAATAGTAGAGAAAACGAGAAATTC
CAATTTTTTACATTGTGTTTCGAAAAAATATACAGGGGTGCTGGTGCTAC
35 TCTGGGCACCAAAGTGGAACCGCCCCTGCACACACACACATAGAGGGA
GAGAGAGAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT
TGGGATGTGATACTTCTTTTGGGAAAATGGAGCAATATCTTTAATATTGT
ATTTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATAACTTTTA
GTTTTTTTTTATTTAATCTGTATATTTAATCATTTTCAGTTTATAAGTTTT
40 ATTTATTTTGGTATACCAGAAAAAAAAGTCTTTTATGTGTTGGATTTAAC
ATAAAAATCTAACAATATTAATCAAAAAGACCAAACATGTGGACAATTAT
GTATATAATTAATTCTCAATGGTCTTAGTGTAACGATATAAAATTTCAAAA
CAATTTTTTCACATTAATAAAAAAACACTTTCAGTCATAATTGTTATAAATTA

TCATTGTATCACAAAATCAGTTCATAACATCACATCCCAAGATCAATAAA
GTGTAAATACTCCTCATGTGTGTACTAATCAAGCCGACGCCTTCCCGCGA
TTCTCACTGGTACCTGAAACACGTAACATAACAACCTGTAAGCATAAAATGC
TTAGTGAGTTCCCCAAAATACCACATACCACATATATGCCTTTCAGGCC
5 ATA ACTCTGTAGGATCTTCCGACCCAAGTGTCTCAGGGGACTTCCGTCCC
GAATCCCGGTAGACCTTCCGGTCCTACCCGATTGACCTTCCGGTCCGTA
TCATACATAACATAACATAACACATACATATCACATAACAACATATAGCAC
ATACATCTCATAACATAAAAGACCTTCCGGTCACATAAAGGTACCCTTCC
AGGTACAGTATAGTGAGAANACTCACCTCGTATGATGTCTAATACCTCAC
10 GTGCTCGATATCCCTGAATCTCGAAACAATGACCTAGCCCCGCCTACTCA
CATAAAGTAATTATTTCAAATCATTACGGCTCTCAAGGCTAGACTACAT
CCCTTCTATAAATCCACAGAAGGGTAAAAGACCATTTTACCCCTCCTTG
ACCCAAAAGTCCAAATGTTGATCAAAACCCCAAAAGTCAACGAAAGACAA
TGGTCAACTTTGACCCTACTCGTGGAGTGCACAAAGGTGACTCGGCAAGT
15 ACATGCGGGTCCTCTGAATCCTTTCAGTCTCTCTTGGCTCGTCGAGTCTT
TCTTCCACCCGACGAGTTACACCTGTCATGAATCGCGGGGCAACCCCGAC
TCGACTTGTCGAGTCCGCTCATGGACTCAACGAGTTCATTCCATGCTCAC
ACTCAAATGACCTCCTGAGGTCAGATCTGTTCTCTAATCCATAGATCTG
ACCTTCCCAAGCTCAATAAACACGTAAAGGTTTGAACCTTGATACTCATGC
20 AACGTCCAAATGATTCTACTTGATGATTTAGCCCCAAATACAACATCCTA
AGTCCATACGACCTTATTTTTCTCAAATAACAACACATATATTTAATTAC
CAATGACAGTAATAGATATCATATAAAGTATTTGTAACACTTTGTAAGAA
CCTTGCTACTATAGGTAAAAAGAAACATTTCAAAGTACATGCCCTAATTA
GAAAAAAAGTTATAAAAAAATAATGACTAGGGGCGTGTTTTTTTACTAG
25 TTTGTATCAAATTATATCAAAATTTAAGGTGGAAAAAGATGACGACCACA
TTAACCAGAAATGTAATTATTTTTTTTATTTGGTAATTTTTTAATATTTGTT
GTGATCTATGTATTTAAAAGTAAATATCAAACAAGAACATAATCCAAACC
CTAAATTGCAAGTCTCGCCCAATTTCTCTATCACTAGTCCTCACTTACGA
TGGCGTTACGTCGCTCTCTCACTTCCTACAACCCATTGTTGCTACTAATT
30 AACTAACGAAAAGTTGAATATCCATATATTTATTTGGATGTGAAATTGA
ACGAATCTCGTCAAATTTTTTATTTTGTGATGGATTTGAGTGGAAGTTT
AGGCAGAACGGGAATGATGGTCTGCAAGTGGTTATAAACATGGGTGAAGA
TAAATGGAGTTGTGCGCGTTGTATTATAGATCTCTTAGGGGTTTGATTCT
TGAGTTATTACTGTATACGTAGCCTCTTTACAACGACCATTCTTCCAAGT
35 ACCATTTGATCTTTTTAGAAATCCAGTTGTCTGAAACACCCTGATTTGGAT
CAAATATCACCAACAACCTCTTAAGAACTGGACTAATTAATTGTTTTCTTG
ATCTTGATAACAAGAGGAAACACGTCACCATATCTTTTATTTTAAATTTG
CTTTTGGTGTATTTTCTTTCTTCCCATTTCTTTCTTGATCTGTTCCAGAT
GGTATTTGGTGTGGATAATTTACACCTGGAGATTGTGAACGATGGGAAGG
40 GGTATGTGATTTACAGAGGATGTGGCTTGTGGTTGAGGATGGTTTATGGC
TGGCCGAGTCTAATTTATATTTATATAAACAATAAATATATAAAACAAG
GGTAAAATATGTATTTAAGCGTCCTCTTTAATGGTGACAATTTTTTACAG
TTTACTCTCTTTGTTTTTTAATTGTGATGCCACGATCGAACTCATTCAT

CCCCCCCCCTTTTTTTTTTAAAAATAAAAAATTAAGAAGGGGTACCACCAT
ATACCCGTGTCAGCTTCTTATTCCCAAGCAGTCAAATAGGGACTTAGGTT
GTATGGAAACAGTTCCGTGACTTGGATGGCAGATAAAATTTAGTAAACTTA
ACCCTTCAATTAACCTACCTTTTTCTTATTAACCTCAATTTCAAGCTAAAT
5 TCTGATTCTTGTTTGAAAATAAGTTGCATCTTTATTTTTGCATATTATCT
TGTTGCATAGGATCCTTAGCATCTTTTAATAGTTTATTTGAAGCTGAAAG
ATCCAACCTAGTTTTGATCTGTTGGCATTTCATCATTTGCAACTGTTTC
TTGAAAAAAAATACCTAAAATCAAATAAACCATTTTCAAATCCAAAATTA
TAAGAGAGAATTGTTAATGGACGTGGAATCATAAATCATTAACACAGTTC
10 AGTACACAAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTAT
CAGAGAAAAGAGACATTACAAGAAGTCACTGATACTAATATTTCTAATGAT
GTTGTATTATTCCCATCCTGTCTCATGCACTCTTTTCATAACCTCCATAA
ACTTAAATTGGAATAATTATGAAGGAGTGGAGGTGGTGTGTTGAGATAGAGA
GTGAGAGTCCAACATGTAGAGAATTGGTAACAACCTACAATAACCAACAA
15 CAGCCTATTATACTTCCCAACCTCCAGGAATTGTATCTAAGGAATATGGA
CAACACGAGTCATGTGTGGAAGTGCAGCAACTGGAATAAATTCTTCACTC
TTCCAAAACAACAATCAGAATCACCATTCCACAACCTCACAACCATAGAA
ATGAGATGGTGTGTCATGGCTTAGGTACTTGTTCGCCTCTCATGGCAGA
ACTTCTTTCCAACCTAAAGAAAGTCAAGATACTTGGGTGTGATGGTATTG
20 AAGAAGTTGTTTCAAACAGAGATGATGAGGATGAAGAAATGACTACATTT
ACATCTACCCACACAACCACCAACTTGTTCCCTCATCTTGATTCTCTCAC
TCTAAAATACATGCACTGTCTGAAGTGTATTGGTGGAGGTGGTGCCAAGG
ATGAGGGGAGCAATGAAATATCTTTCAATAATACCACTACAACCTACCGAT
CAATTTAAGGTATGTTTGTACATATTTAATTATATATTTAATTTCCCTGT
25 TAATTTCCCTTTTCTTTGCAATATTCTATGCGAACTCAAGAATGGGATTG
GAGGCATATAAAGTTACATTCATTTGAACAAGTATTACCTTTTATTTGTT
ATTTATCATTTTCATATCAAGTACCTATAACATTTCTTTTTTATTTTTCT
AATTAGAAGAGGTCCACATGTCTAATTAGGTTTTCCATTCTATGTGTAAC
CTCTATTCTCTCTGTAATCAAGCATCTTAGATTATTTATCCATTTTCATA
30 ATTGTTGTTTATTTTTACAGTTTTTTTTTTTTTATTTAATTTTAATAATTTAA
TTTTAATTTATTTATTTTTTTTTTTTTGGTAATTGCAACCTGTCATATAT
TCAAGTCTTAATGTAACATAATAATACATTTTATACCCACTATACTAAGA
TAATAATTACCTAAAGGGATGGATGCCATGCACTGCTACACTTCAGNAA
CTCTAGTAAGGGCAGTTATGGAAGTTCAATAAAAATGATAATGGCATCTTT
35 TGATGGGTAATATAGGCAATTTAAGTTTTATTTCTGTAAAGCAGTATTT
AGCTAGTAGTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTCAAATCT
GGTCATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATTAGGGG
TCATCAGGTGACAGATATTGTAGAATAGAACAAATATGTAATATTACCCAA
AACTATTTTTTCTAAGGTTGCTCTGTAAATATGTGCTTCTTGATTTC
40 TTGAATTTGCATTCGTATATTTTAGGTGGTAAACTGATTGTCTCTTCAAT
AAATCCTGAAATTAATTAACAAAAAACAACAAAGTACATTTTTGATTT
GGAGAGCACTGGTATCATTTAGTATAGAAAAAACTAGATTTTGAATTAY
CTTTCTTATATAAAAGTTGTGTATATAGTTAATTAGTTTTACATCATTT

TTCTATGTGTTGTTGCAGTTGTCTGAAGCAGGTGGTGTGTTGTTGGAGCTT
ATGCCAATACTCTAGAGAGATAGAGATATATAGGTGTGATGCACTGTCAA
GTGTAATTCCATGTTACGCAGCAGGACAAATGCAAAAGCTGCAAGTGCTG
ACAGTCAGTTCTTGTAATGGTCTGAAGGAGGTATTTGAACTCAATTAGG
5 GACGAGCAGCAACAAAAACAACGAGAAGAGTGGTTGTGAGGAAGGAATTC
CAAGAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTGGAA
ATCTACGGTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGA
AAGCCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTACTACTCTTGTG
AATCTTCCAAACCTCAAAGAAATGAGGTTGGAGTGGCTAAGTAATCTGAG
10 GTATATATGGAAGAGCAATCAGTGGACAGCATTTGAGTTTCCAAACCTAA
CAAGAGTTGAAATTTGTGAATGTAATTCATTAGAACATGTATTTACTAGT
TCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTACATATATTTAACTG
CAGTCTGATGGAGGAGGTAATTGTTAAGGATGCAGATGTTTCTGTAGAAG
AAGACAAAGAGAAAGAATCTGATGGCAAGACGAATAAGGAGATACTTGTG
15 TTACCTCATCTAAAGTCCTTGAAATTACAACCTCTTCGAAGTCTTAAGGG
GTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTATTGGATACTTTAG
AAATCAAAAGATGCCCAACAATAACCACCTTCACCAAAGGAAATTCGCT
ACTCCACAACATAAAAGAAATACAAACAAATTTTGGCTTCTTTTATGCTGC
AGGGGAAAAAGACATCAACTCTCTTATAAAGATCAAACAACAGGTAAATC
20 AGATCTTTGTTGCTTTAATAATTCTTAAACTACATTTGAAAAGCTTCATG
CAAGTTTTTTTGTATATTGTCAAAAACCGCAACCTACATTCAGCTTTAT
ATTTATGTACTTTATGCAGGATTTCAAACAAGACTCAGATTAATGTGAAG
TGAATATTAAAGGTAAATTATATTTTCATGTTCTAGTTGCCTATTAATT
AATGGCCTTTTAGTTCATGATTTTTGGATGTATTCTTCATGATGATGTGA
25 ATCTTCTAATACCCCATTCATTGTTTGGTTGAATGTTGACTCTATGTCAG
GATGAATATTCAAGGGAAGAATTGTTTCATCAWATGAAGGACATTAAAGAA
CATGGATGCTATGAAGATGTTGGGAAAACATATGTATCAAGTGGCAARCT
GCTTAATGATCTAAGTTTGTGGTTGANGATGTTGATTTTAATATTTCAA
ATTCATTGGTTATATGGGCTTATCAATAGTGTTAATGGGATAATGAGTGA
30 CTTAACCTAAATTATGTTGTTGGTAAATGTTGGACAAGTATGGAAAATTA
GGAATGACTTGTGAAAAAAAAAATAAAAAAAAAA (SEQ ID NO:94)

RG2D deduced polypeptide sequence (SEQ ID NO:95)

MAMETANEIKQVVPVLMVPINDYLRYVVSCKYISDMDLKMKEKEDNVEE
35 HKNHNISNRLEVPAAQVQSWLEDVEKINAKVETVPKDVGCCFNLKIRYRAGDAF
NIIIEIDSVMRRHSLITWTDHPIPLGRVDSVMASSTLSTEHNDFQSREVRFSEALKA
LEANHMIALCGMGRVGKTHMMQRLKKVAKEKRKFGYIIEAVIGEISDPIAIQVVVA
DYLIELKESDKKTRAELRQGFKAQSDGGNTKFLIILDDVWQSVLEDIGLSPSPN
QGVDFKVLLTSRDEHVCVSMGVEANSIINVGLLIEAEAQRLFQQFVETSEPELHKIG
40 EDIVRRCCGLPIAIKTMACTLRNKRKDAWKDALSRHQHHDIGNVATAVFRTSYENL
PDKETKSVFLMCGLFPEDFNIPTEELMRYGWGLKLFDRVYTHIARNRLNTCIERLV
QANLLIGSDNGVHVKMHDLVRAFVLGMYSEVEQASIVNHGNMGPWPENDMIVH

SCKRISLTCKGMIEIPVDLKF PKLTILKLMHGD KSLKFPQEFYEGMEKLQVISYDKM .
KYPLLPLAPQCSTNIRVLHLTECSLKMFD CSSIGNLSNLEVL SFANSRIEWLPSTVRN
LKKLRLLDLRFCDGLRIEQGV LKSLVKLEEFYIGNA YGFIDDNCKDMAERSYNLSA
LEFAFFNKA EVKNMSFENLERFKISV GCSFDGNISMSSH SYENMLQLVTNKGDVL
5 DSKLNGFLFKTEVLF LSVHGMNDLEDVEVKSTHPTQSSSFCNLKVRIISKCVELRYL
FKLHVANTLSSLEHLEVCGCENMEELIHTGIGGCGEETITFPKLKSLSLSQLPKLSGL
CHNVNII GLPHLVDLKLKGIPGFTVIYPQNKLR TSSLLKEEVVIPKLETLOIDGMENL
EEIWPCELSGGEKVKLREIKVSSCDKLVNLFPHNPMSLLHHLEELKVKNCRSIESLF
NIDLDCVSAIGEEDNKSILRRIKVKNLGK LREVWRIKGADNSRPLIHGFPAVESISIW
10 GCKRFRNIFTPITANFDL VALLEIHIGNYRENHESEEQIEILSEKETLOEVTDTNISND
VVLFPSCLMHSFHNHLHKLKLENYEGVEV VFEIESESPTCRELVTTHNNQQQPILPN
LQELYLRNMDNTSHVWKCSNWNKFFTL PKQQSESPFHNLT TIEMRWCHGFRYLF
PLMAELLSNLKKVKILGCDGIEEVVSNR DDEDEEMTTFTSTHTTTNLFPHLDSLTLK
YMHCLKCIGGGGAKDEGSNEISFNNTTTT DQFKLSEAGGVCWSLCQYSREIEIYRC
15 DALSSVIPCYAAGQMQLQVLT VSSCNGLKEVFETQLGTSSNKNNEKSGCEE GIPR
VNNNVIMLPNLKILEIYGCGGLEHIFTF SALESRLQLQELTIKGYT L VNLPNLKEM
RLEWLSNLRYIWKSNQWTA FEFPNLTRVEICECNSLEHVFTSSMVGSLQLQELHIF
NCSLMEEVIVKDADVSVEEDKEKESD GKTNKEILVPLHLKSLKLQLLRS LKGFSLGK
EDFSFPLD TLEIKRCPTITTTFTK GNSATPQLKEIQTNFGFFYAAGEKDINS LKIKQQ
20 DFKQDSD.CEVNIK

RG2E polynucleotide sequence (SEQ ID NO:96)

TGGGAAGACACAATGATGCAAAGGTTGAAGAAGGTTGCTAAAGAAAATAGAAT
GTTCAATTATATGGTTGAGGCAGTTATAGGGGAAAAGACAGACCCACTTGCTAT
25 TCAACAAGCTGTAGCGGATTACCTTTGTATAGAGTTAAAAGAAAGCACTAAACC
AGCAAGAGCTGATAAGCTTCGTGAATGGTTTAAGGCCAACTCTGGAGAAGGTA
AGAATAAGTTCCTTGTAATATTTGATGATGTTTGGCAGTCCGTTGATCTGGAAG
ACATTGGTTTAAGTCATTTTCCAAATCAAGGTGTCGACTTCAAGGTCTTGTTGA
CTTCACGAGACGAACATGTTTGCACAGTAATGGGGGTTGAAGCTAATTCAATTC
30 TTAATGTGGGACTTCTAGTAGAAGCAGAAGCACAAAGTTTGTTCAGCAATTTG
TAGAAACTTTTGAGCCCGAGCTCCATAAGATAGGAGAAGATATCGTAAGGAAG
TGTTGTGGTTTACCTATTGCCATTA AAACCATGGCATGTACTCTAAGAAATAAA
AGAAAGGATGCATGGAAGGATGCACTTTTGCATTTAGAGTACCATGACATTAGC
AGTGTTGCGCCCAAAGTCTTTGAAACGAGCTACCATAATCTCCACAACAAGGAG
35 ACTAAATCTGTGTTTTTGATGTGTGGTTTTTTTCTGAAGACTTCAATATTCCAA
TCGAGGAGTTGATGAGGTATGGATGGGGCTTAAAGATATTTGATAGAGTTTATA
CTATTAGACAAGCAAGAATCAGGCTCAACACCTGCATTGAGCGACTGGTGCAG
ACAAATTTGTTAATAGAAAGTGATGATGGTGTGCACGTCAAGATGCATGATCTG
GTCCGTGCTTTCGTTTTGGTTATGTTTTCTGAAGTTGAACATGCTTCAATTATCA
40 ACCATGGTAATATGCTTGGATGGCCTGAAAATTATATGACCAACTCTTGCAAAA
CAATTTTCATTAACATGCAAGAGTATGTCTGAATTTCCGGGAGATCTCAAGTTTC
CAAACCTAACGATTTTGAAACTCATGCATGGAGATAAGTTGCTAAGATATCCTC

AAGACTTTTATGAAGGAATGGAAAAGCTCTGGGTTATATCATATGATGAAATGA
AGTATCCATTGCTTCCCTCGTTACCTCAATGCTCCATCAACCTTCGAGTGCTTCA
CCTCCATCGATGCTCATTAATGATGTTTGATTGCTCTTGTATTGGAAATATGTTG
AATCTGGAAGTGCTTAGCTTTGTAAATCTGGCATTGAATGGTTACCTTCCACA
5 ATAGGAAATTTAAAGAAGCTAAGGTTACTTGATCTGAGAGATTGTTATGGTCTT
CGTATAGAAAAAGGTGTCTTGAAAAATTTGGTGAAAATTGGAGGAATTTATATT
GGTAGAGCAGATATTTTATAGAT

RG2E deduced polypeptide sequence (SEQ ID NO:97)

10 WEDTMMQRLKKVAKENRMFNYMVEAVIGEKTDLAIQQA VADYLCIELKESTKP
ARADKLREWFKANS GEGKNKFLVIFDDVWQSV DLEDIGLSHFPNQGVDFKVLLTS
RDEHVCTVMGVEANSILNVGLLVEAEAQSLFQQFVETFEPELHKIGEDIVRKCCGL
PIAKTMACTLRNKRKDAWKDALLHLEYHDISSVAPKVFETSYHNLHNKETKSVFL
MCGFFPEDFNIPIEELMRYGWGLKIFDRVYTIRQARIRLNTCIERLVQTNLLIESDDG
15 VHVKMHDLVRAFVLVMFSEVEHASIINHGNMLGWPENYMTNSCKTISLTCKSMSE
FPGDLKFPNL TILKLMHGDKLLRYPQDFYEGMEKLWVISYDEMKYPLLPSLPQCSI
NLRVLHLHRCSLMMFDCSCIGNMLNLEVL SFVKSGIEWLPSTIGNLKKLRLLDLRD
CYGLRIEKGVLKNLVKIGGIYIGRADIL.

RG2F polynucleotide sequence (SEQ ID NO:98)

CTGTGGAAGACACAATGATGCAAAGGCTGAAAAAGGTTGTGCATGAAAAGAAA
ATGTTTAACTTTATTGTTGAAGCAGTTATAGGGGAAAAGACAGACCCCGTTGCC
ATTCAGGATGCTATAGCAGATTACCTAGGTGTAGAGCTCAATGAAAAATCTAAG
CAAGCAAGAGCTGATAAGCTCCGTCAAGGATTCAAGGACAAATCAGATGGAGG
25 CAAAATAAGTTCTTTGTAATACTTGACGATGTTTGGCAGTCTGTTGATCTGGA
AGATATTGGTTTAAAGTCCTTTTCCAAATCAAGGCGTCGACTTCAAGGTCTTGTT
GACATCACGAGACAGACATGTTTGCACAGTGATGGGGGTGAAGCCAAATTAA
TTCTAAACGTGGGACTTCTAATTGAAGCTGAAGCACAAAGTTTGTTCACCAAT
TTGTTGTCACTTCTGAGCCCGAGCTCCATAAGATAGGAGAAGATATTGTAAAGA
30 AGTGTTCGGTCTGCCAATTGCCATCAAAACCATGGCATGTACTCTACGACATA
AAAGAAAGGATGCATGGAAGGATGCATTTACGTTTAGAGCACCATGACATT
CAAAGTGTTGTGCCTAAAGTATTTGAAACGAGCTACAACAATCTCAAAGACAA
GGAGACTAAATCCGTATTTTTGATGTGTGGTTTGTTCCTGAAGACTTGGATAT
ACCTATCGAGGAGTTGATGAGGTATGGATGGGGCTTAAGATTATTTGATAGAGT
35 TAATACTATTACACAAGCAAGAAACAGGCTCAACACCTGCATTGAGCGACTGG
TGCACACAAATTTGTTAATTGAAAGTGTTGATGGTGTGCATGTCAAGATGCATG
ATCTGGTTCGTGCTTTTGT TTTTGGGAATGTTTTCTGAAGTGGAGCATGCTTCAAT
TGTCAACCATGGTAATATGCCCCGAGTGGACTGAAAATGATATGACTGACTCTTG
CAAACAAATTTCAATTAACATGCAAGAGTATGTTGGAGTTTCTGGAGACCTCAA
40 GTTTCCAAACCTAAAGATTTTGAAACTTATGCATGGAGGTAAGTCACTAAGGTA
TCCTCAAGACTTTTATCAAGGAATGGAAAAGCTGGAGGTTATATCATACGATGA
AATGAAGTATCCATTGCTTCCCTCGTTGCCTCAATGTTCCACCATCCTTCGAGTG

CTTCATCTCCATGAATGTTCAATTAAGGATGTTTGATTGCTCTTCAATCGGTAATC.
TTTTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAGCATTGAATTGTTACCTTC
CGTAATTGGAATTTGAAGAAGTTGCGGCTGCTAGATTTGACAACTGTTATGG
TGTTTCGTATAGAAAAGGATGTCTTGAAAAATTTGGTGAAACTTGAAGAGCTTTA
5 TATTAGGAATGGTCTACCAGTTTACAGAGGAT

RG2F deduced polypeptide sequence (SEQ ID NO:99)

VEDTMMQRLKKVVHEKKMFNFIVEAVIGEKTDPVAIQDAIADYLGVELNEKSKQA
RADKL RQGFKDKSDGGKNKFFVILDDVWQSV DLEDIGLSPFPNQGVDFKVLLTSRD
10 RHVCTVMGVEAKLILNVGLLIEAEAQSLFHQFVVTSEPELHKIGEDIVKKCFGLPIAI
KTMACTLRHKRKDAWKDALSRLEHDIQSVVPKVFETSYNNLKDKETKSVFLMCG
LPEDLDIPIEELMRYGWGLRLFDRVNTITQARNRLNTCIERLVHTNLLIESVDGVH
VKMHDLVRAFVLGMFSEVEHASIVNHGNMPEWTENDMTDSCKQISLTCKSMLEFP
GDLKFPNLKILKLMHGGKSLRYPQDFYQGMKLEVISYDEMKYPLLPQLPQCSTILR
15 VLHLHECSLRMFDCCSIGNLNFMEVLSFANSSIPELLPSVIGNLKKLRLDLTNCYGV
RIEKDVLKNLVKLEELYIRNGLPVYRG

RG2G polynucleotide sequence (SEQ ID NO:100)

GAAGACACGATGATGAAGAACTGAAGGAGGTCGTGGGACAAAAGAAATCATT
20 AATATTATTATTCAAGTGGTCATAGGAGAGAAGACAAACCCTATTGCAATTGAG
CAAGCTGTAGCAGATTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGC
AAGAGCTGATAAGCTTCGTAAACGGTTTGAAGCCGATGGAGGAAAGAATAAGT
TCCTTGTAATACTTGACGATGTATGGCAGTTTGTTCGATCTTGAAGATATTGGTTT
AAGTCCTCTGCCAAATAAAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGA
25 TTCACATGTTTGCACTCTGATGGGAGCTGAAGCAAATTCAATTCTTAATATAAA
AGTTTTTAAAGATGTAGAAGGACAAAGTTTGTTCGCCAGTTTGCTAAAAATGC
GGGTGATGATGACCTGGATCCTGCTTCAATGGGATAGCAGATAGTATTGCAAG
TAGATGTCAAGGTTTGCCCATTCATCAAAACCATTCCTTAAGTCTTAAAGG
TAGAAGCAAGTCTGCATGGGACGTTGCACTTTCTCGTCTGGAGAATCATAAGAT
30 TGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAATTAGCTACGACAATCTCCA
AGATGAGGTTACTAAATCTATTTTTTTACTTTGTGCTTTATTCCTGAAGATTTT
GATATTCCTACTGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATA
GAAGCAAAAACCTATAAGAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCG
GCTTAGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGTGTCAAGAT
35 GCACGATGTGGTGCGTGATTTTGTGTTTGCATATATTCTCAGAAGTCCAACACGC
TTCAATTGTCAACCATGGTAACGTGTCAGAGTGGCTAGAGGAAAATCATAGCAT
CTACTCTTGTAAGAATTTCAATTAACATGCAAGGGTATGTCTCAGTTTCCCAA
AGACCTCAAATTTCCAAACCTTTCAATTTTGAACTTATGCATGGAGATAAGTC
ACTGAGCTTTTCTGAAAACCTTTTATGGAAAAGATGGAAAAGGTTCAAGTAATATC
40 ATATGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACCAA
CGTTCGAGTGCTTCATCTTCATTACTGTTCAATTAAGGATGTTTGATTGCTCTTCA
ATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAACATTGAA

TGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACA
AATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTAAAAAATTTGGTCAAACCTT
GAAGAGCTTTATATGGGTGTTAATCGTCCGTATGGACAGGCCGTTAGCTTGACA
GATGAAAA

5

RG2G deduced polypeptide sequence (SEQ ID NO:101)

RHDDEELKEVVVGQKKSFNIIQVVIGEKTNPQIAIQAVADYLSIELKENTKEARADKL
RKRFEADGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTSRDSHVCTL
MGAEANSILNIKVLKDVEGQSLFRQFAKNAGDDDLDPAFNGIADSIASRCQGLPIAI
10 KTIALSLKGRSKSAWDVALSRLENHKIGSEEVVREVFKISYDNLQDEVTKSIFLLCAL
FPEDFDIPTTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVK
MHDVVRDFVLHIFSEVQHASIVNHGNVSEWLEENHSIYCKRISLTCKGMSQFPKDL
KFPNLSILKLMHGDKSLSPENFYGKMEKVQVISYDKLMYPLLPSLECSTNVRVLH
LHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRID
15 NGVLKNLVKLEELYMGVNRPYQAVSLTDE

RG2H polynucleotide sequence (SEQ ID NO:102)

TGAAGGAGGTTGTGGAACGAAAGAAAATGTTCAGTATTATTGTTCAAGTG
GTCATAGGAGAGAAGACAAACCTATTGCTATTCAGCAAGCTGTAGCAGA
20 TTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATA
AGCTTCGTAAATGGTTTCGAGGCCGATGGAGGAAAGAATAAGTTCCTTGTA
ATACTTGACGATGTATGGCAGTTTGTGCGATCTTGAAGATATTGGTTTAAG
TCCTCTGCCAAATAAAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAG
ATTCACATGTTTGCCTCTGATGGGAGCCGAAGCCAATTCAATTCTCAAT
25 ATAAAAGTTTTTAACAGCTGTAGAAGGACAAAGTTTTGTTCCGCCAGTTTGC
TAAAAATGCGGGTGATGATGACCTGGATCCTGCTTCAATAGGATAGCAG
ATAGTATTGCAAGTAGATGTCAAGGTTTGCCCATTGCCATCAAAACCATT
GCCTTAAGTCTTAAAGGTAGAAGCAAGCCTGCGTGGGACCATGCGCTTTC
TCGTTTGGAGAACCATAAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTT
30 TTAAAATTAGCTATGACAATCTCCAAGATGAGATTACTAAATCTATTTTT
TTACTTTGTGCTTTATTTCTGAAGATTTTGATATTCCTACTGAGGAGTT
GATGAGGTATGGATGGGGCTTGAAATTATTTATAGAAGCAAAAACCTATAA
GAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCGGCTTAGGGAGACA
AATTTGTTATTTGGAAGCGATGACATTGGATGCGTCAAGATGCACGATGT
35 GGTGCGTGATTTTGTGTTTGCATATATTCTCAGAAGTCCAGCACGCTTCAA
TTGTCAACCATGGTAACGTGTGAGAGTGGCTAGAGGAAAATCATAGCATC
TACTCTTGTAAGAATTTTCAATTAACATGCAAGGGTATGTCTGAGTTTCC
CAAAGACCTCAAATTTCAAACCTTTCAATTTTGAACTTATGCATGGAG
ATAAGTCGCTGAGCTTTCCTGAAAACCTTTTATGGAAGATGGAAGAGGTT
40 CAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACT
TGAATGCTCCACTAACGTTTCGAGTGCTTCATCTCCATTATTGTTCAATTA
GGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTC

AGCTTTGCTAATTCTAACATTGAATGGTTACCATCTACAATTGGAAATTT
GAAGAAGCTAAGGCTACTAGATTTGACAAATTGTAAAGGTCTTCGTATAG
ATAATGGTGTCTTAAAAAATTTGGTCAAACCTGAAGAGCTTTATATGGGT
GTTAATCATCCGTATGGAC

5

RG2H deduced polypeptide sequence (SEQ ID NO:103)

KEVVERKKMFSIIVQVVIGEKTNPQIAIQQAVADYLSIELKENTKEARADKLRKWFEA
DGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTSRDSHVCTLMGAEAN
SILNIKVLTAVEGQSLFRQFAKNAGDDDLDPAFNRIADSIASRCQGLPIAKTIALSLK
10 GRSKPAWDHALSRLENHKIGSEEVVREVFKISYDNLQDEITKSIFLLCALFPEDFDIP
TEELMRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVVR
DFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEFPKDLKFPNLSI
LKL.MHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNVRVLHLHYCSL
RMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRIDNGVLKN
15 LVKLEELYMGVNHPYG

RG2I polynucleotide sequence (SEQ ID NO:104)

AAGAAGAGCTGAAGGAGGTTGTGGAACAAAAGAAAACGTTCAATATTATT
GTTCAAGTGGTCATAGGAGAGAAGACAAACCCTATTGCTATTCAGCAAGC
20 TGTAGCAGATTCCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAA
GAGCTGATAAGCTTCGTAAATGGTTCGAGGCTGATGGAGGAAAGAATAAG
TTCCTCGTNATACTTGACGATGTATGGCNGTTTGTGATCTTGAAGATAT
TGGTTTAAGTCCTCATCCAAATAAAGGTGTCANCTTCAAGGTCTTGTTGA
CGTCAAGAGATTACATGTTTGCACCTCTGATGGGAGCTGAAGCCAATTCA
25 ATTCTCAATATAAAAGTTTTAAAGATGTAGAAGGAAAAAGTTTGTCCG
CCAGTTTGCTAAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCATTG
GGATAGCAGATAGTATTGCAAGTAGATGTCAAGGTTTGCCCATTTGCCATC
AAAACCATTGCCTTAAGTCTTAAAGGTAGAAGCAAGTCTGCATGGGACGT
TGC.ACTTTCTCGTCTGGAGAATCATAAGATTGGTAGTGAAGAAGTTGTGC
30 GTGAAGTTTTTAAATTAGCTATGACAATCTCCAAGATGAGGTTACTAAA
TCT.ATTTTTTTACTTTGTGCTTTATTTCTGAAGATTTTGATATTCCTAC
TGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATAGAAGCAA
AACTATAAGAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCGGCTT
AGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGCGTCAAGAT
35 GCACGATGTGGTGCGTGATTTTGTGTTTGCATATATTCTCAGAAGTCCAGC
ACGCTTCAATTGTCAACCATGGTAATGTGTCAGAGTGGCTAGAGGAAAAT
CATAGCATCTACTCTTGTAAGAAGATTTTCAATTAACATGCAAGGGTATGTC
TGAGTTTCCCAAAGACCTCAAATTTCCAAACCTTTCAATTTTGAACTTA
TGC.ATGGAGATAAGTCGCTGAGCTTTCCTGAAAACCTTTATGGAAAGATG
40 GAAAAGGTTCAAGTAATATCATATGATAAATTGATGTATCCATTGCTTCC
CTC.ATCACTTGAATGCTCCACCAACCTTCGAGTGCTTCATCTCCATGAAT
GTTCAATTAAGGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATG

GAAGTGCTCAGCTTTGCTAATTCTGGCATTGAATGGTTACCATCTACAAT
TGGAAATTTGAAGAAGCTAAGGCTACTGGATCTGACAGATTGTGGAGGTC
TTCATATAGATAATGGCGTCTTAAAAAATTTGGTCAAACCTGAAGAGCTT
TATATGGGTGCTAATCGTCTGTTTGGAAAGTGCCAT

5

RG2I deduced polypeptide sequence (SEQ ID NO:105)

EELKEVVEQKKTFNIIQVQVIGEKTNPIAIQQA VADSLSELKENTKEARADKLRKWF
EADGGKNKFLVILDDVW?FVDLEDIGLSPHPNKG V?FKVLLTSRDSHVCTLMGAEA
NSILNIKVLKDVEGKSLFRQFAKNAGDDDLDPAFIGIADSIASRCQGLPIAKTIALSL
10 KGRSKSAWDVALSRLENHKIGSEEVVREVFKISYDNLQDEVTKSIFLLCALFPEDFDI
PTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVV
RDFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEFPKDLKFPNLS
ILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNLRVLHLHECSL
RMFDCSSIGNLLNMEVLSFANSIGIEWLPSTIGNLKKLRLLDLTDCGGLHIDNGVLKN
15 LVKLEELYMGANRLFGKCH

RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107)

ATGTCCGACCCAACAGGGATTGTTGGTGCCATTATTAACCCAATTGCTCA
AACGGCCTTGGTTCCCCTTACAGACCATGTAGGCTACATGATTTCTCTGCA
20 GAA AATATGTGAGGGACATGCAAATGAAAATGACAGAGTTAAATACCTCA
AGAATCAGTGCAGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCA
GATTCATCTCAAATTAAGGATTGGTTGGACCAAGTAGAAGGGATCAGAG
CGAATGTTGCAAACCTTTCCAATTGATGTCATCAGTTGTTGTAGTCTCAGG
ATCAGGCACAAGCTTGGACAGAAAGCCTTCAAGATAACTGAGCAGATCGA
25 AAGTCTAACGAGACAAAATTCGCTGATTATCTGGACTGATGAACCTGTTC
CCCTGGGAAGAGTTGGTTCCATGATTGCATCCACCTCTGCAGCATCAAGT
GATCATCATGATGTCTTCCCTTCAAGAGAGCAAATTTTTAGGAAAGCACT
AGAAGCACTTGAACCCGTCCAAAAATCCACATAATAGCCTTATGGGGGA
TGGGCGGAGTGGGGAAGACCACGATGATGAAGAAGCTGAAAGAGGTCGTG
30 GAACAAAAGAAAACGTGCAATATTATTGTTCAAGTGGTCATAGGAGAGAA
GACAAACCCTATTGCTATCCAGCAAGCTGTAGCAGATTACCTCTCTATAG
AGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTAAACGG
TTCGAAGCCGATGGAGGAAAGAATAAGTTCCTTGTAATACTTGACGATGT
ATGGCAGTTTTTCGATCTTGAAGATATTGGTTTAAGTCCTCTGCCAAATA
35 AAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGATTCACATGTTTGC
ACTCTGATGGGAGCTGAAGCCAATTCTATTCTCAATATAAAAGTTTTAAA
AGATGTAGAAGGAAAAAGTTTTGTTCCGCCAGTTTGCTAAAAATGCGGGTG
ATGATGACCTGGATCCTGCTTTCATTGGGATAGCAGATAGTATTGCAAGT
AGATGTCAAGGTTTGGCCATTGCCATCAAAACCATTGCCTTAAGTCTTAA
40 AGGTAGAAGCAAGTCTGCATGGGACGTCGCACTTTCTCGTCTGGAGAATC
ATAAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAATTAGCTAT
GACAATCTCCAAGATGAGGTTACTAAATCTATTTTTTTACTCTGTGCTTT

ATTTCTGAAGATTTTGATATTCCTATTGAGGAGTTGGTGAGGTATGGGT
GGGGCTTGAAATTATTTATAGAAGCAAAAACCTATAAGAGAAGCAAGAAAC
AGGCTCAACAACCTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGG
AAGTCATGACTTTGGGTGCGTCAAGATGCACGATGTGGTGCGTGATTTTG
5 TTTTGCATATGTTTTCAGAAGTCAAGCATGCTTCAATTGTCAACCATGGT
AACATGTCAGAGTGGCCAGAGAAAAATGATACCAGCAACTCTTGTA AAAAG
AATTTCAATTAACATGCAAGGGTATGTCTAAGTTTCCTAAAGACATCAACT
ATCCAAACCTTTTGATTTTGAACTTATGCATGGAGATAAGTCGCTGTGC
TTTCTGAAAACCTTTATGGAAAGATGGAAAAGGTTCAAGGTAATATCATA
10 TGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACTA
ACGTTTCGAGTGCTTCATCTCCATTATTGTTCAATTAAGGATGTTTGATTGC
TCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTC
TAACATTGAATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGC
TACTAGATTTGACAAATTGTAAAGGCTTTCGTATAGATAATGGTGTCTTA
15 AAAAATTTGGTCAAACCTGAAGAGCTTTATATGGGTGTTAATCGTCCGTA
TGGACAGGCCGTTAGCTTGACAGATGAAAACCTGCAATGAAATGGTAGAAG
GTTCCAAAAA ACTTCTTGCACTAGAATATGAGTTGTTTAAATACAATGCT
CAAGTGAAGAATATATCCTTCGAGAATCTTAAACGATTCAAGATCTCAGT
GGGATGTTCTTTACATGGATCTTTCAGTAAAAGCAGGCACTCATACGAAA
20 ACACGTTGAAGTTGGCCATTGACAAAGGCGAACTATTGGAATCCCGAATG
AACGGGTTGTTTGAGAAAACGGAGGTTCTTTGTTTAAAGTGTGGGGGATAT
GTATCATCTTTCAGATGTTAAGGTGAAGTCCTCTTCGTTCTACAATTTAA
GAGTCCTTGTCGTTTCAGAGTGTGCAGAGTTGAAACACCTCTTCACACTT
GGTGTTGCAAATACTTTGTCAAAGCTTGAGCATCTTAAAGTCTACAAATG
25 CGATAATATGGAAGAACTCATACATACCGGGGGTAGTGAAGGAGATACAA
TTACATTCCCCAAGCTGAAGCTTTTATATTTGCATGGGCTGCCAAACCTA
TTGGGTTTGTGTCTTAATGTCAACGCAATTGAGCTACCAAACTTG TGCA
AATGAAGCTTTACAGCATTCCGGGTTTCAAGCATTTATCCGCGGAACA
AGTTGGAAGCATCTAGTTTGTTGAAAGAAGAGGTACATATACATATAGTT
30 TATGTTAATACATTTTAAACAATCTTTCAACTAAAAGTTTCAGAATATA
TCTGTATTTTGATTGTATGATGTGTTAGTGTGTTGGATGTGGCTATTAAAG
GATAATTATTTGGCAGGTTGTGATTCCTAAGTTGGATATACTTGAAATTC
ATGACATGGAGAATTTAAAGGAAATATGGCCTAGTGAGCTTAGTAGAGGT
GAGAAAGTTAAGTTGAGAAAGATTAAAGTGAGAAATTGTGATAAACTTGT
35 GAATCTATTTCCACACAATCCCATGTCTCTGCTGCATCATCTTGAAGAGC
TTATAGTCGAGAAATGTGGTTCCATTGAAGAGTTGTTCAACATCGACTTG
GATTGTGCCAGTGTAATTGGAGAAGAAGACAACAACAGCAGCTTAAGAAA
CATCAATGTGGAGAATTCAATGAAGCTAAGAGAGGTGTGGAGGATAAAAAG
GTGCAGATAACTCTCGTCCCCTCTTTCGTGGCTTTCAAGTTGTTGAAAAG
40 ATAATCATTACGAGATGTAAGAGGTTTACAAATGTATTCACACCTATCAC
CACAAATTTTGATCTGGGGGCACTTTTGGAGATTTCAAGTTGATTGTAGAG
GAAATGATGAATCAGACCAAAGTAACCAAGAGCAAGAGCAGGTATGGATT
TCAATTTTACTCTTTTACTTAATTAATGATTAAGCCCCTGCTTTTTTAATA

AAAAGGGGACAAACCATTTCTTGACTTAATGTTGCAATACAAGTCATGTA
TAAGAGTGATTAACCTTTTTTTTATTTATAAAATAACTACAAAACATGTTT
TTTCATTATAGATCATGTATAAATGTGACTAATTTTTTTTCATCGCCTAAC
TTTTGTTGATAAATCATTAGAAATGTCACTAATTACTTTTTAGTATTTAT
5 AAAATAACTACAAAACATGTTTTTTTCATTATAGATCATGTATATATCAAC
TAAAAATATTATTCCCTTACACAAAAAAGGTTCAAGAAAGCCTGTA
TTTCGAAATAACTAAAAAGAAATATTTGATATTCATAAGAGAAATTTT
TTTCTAAACATGATCGCAAATGATTAAAACTTAAATTAAACTAAAAAGA
TTTTTATATATGTTATNCAAAATTAATAATTTGAAATTAAGTTTATAATTC
10 TNGTNTCACAAGGGATATATATAGTAAATATTATTTTTTTTGCAGTCAT
GCATAGTTGTATTTTTTAAATGATTTATTAACGTGGTAGGAGTGGAACCA
CTCAATCTAGTAGACCCACTATCACATGTCACATCAGCTTTACATCTATT
TTTCTTTCTCCTTTTTTCATCTTTTTTAAACTCATAACACNTAAAANTANC
ATATTTTCCAACACACTNAACTCATTGTCACATTATTATTTTTTAATTTAA
15 TTAATTTNGAAATTAATAATTAANTAAANCNTAACATTTTTTAATTAATA
AATATTAATCCAAATAAAAAANTNCACGATAAATTAATAANGTTTANTTTG
GAAAAAANCC (SEQ ID NO:106)
Sequence gap
ATAACCCTTTCAAGGGTCAACTCAAGTCCAAGTTAAAGTCAAGGTCAAAA
20 CCTTGGTTAAAGTCAACTTTGGTCAAAGTCAACATCTACTTGACTCACCT
CACCGAGTTGGTCCACCAACTTGTCTGAGTCCCTTAATCCACAACTTCAA
GAACCTCGATCCTACTCGTCGAGTCTTTCAAGAACTCTTCGAGTTTCCAT
TACACAGAATCGGGACCTTTTGCTCATGACTCGCCGAGTTCATCCTTGAA
CTTGTCGAGTCTAGCTTCATACGAGTTCGAGTGTTTAGTCCTTGACTCGT
25 CGAGTTCTTCCTTGAACCTCGTCGAGTCCATCTTCGTATAGTTGGGACATT
GCCCTGAACTCACCGAGTTCATCATTGAACCTCATCGAGTCCCTTCGATCTT
CAAGTCCATAATCCTGTCCATCTTGTTGAGTCCTCTTCTAGACTCAACCA
GATTCCTCAGAAACAGAAAAGGTTAGGGAACCATTACCTGACTCGCCGAG
TCCCAAGAACGAATCCCCGAGTCCCCCAATGTCCATGACCATAACAATCGA
30 TTTTCGTTGGGCTCATTGCATCCAAAGCATAGATCTAACCTCCTAGGGTC
CATATTACACGTAAAGCTACGAACCTTGACGTCCATGCATGGGGGATTTGG
CTCAAATGGCATTAAAATGGGGTTTATCTGATGCATGGGACTCCCATGGC
CATAAAGTTAACACCTTTATGCCATGGGAATCCTCAATGGTTCCATATCT
GAAGTTAACTCTACAATATGTTCTAAACCCGAAGGTGGCTTAGAAATG
35 CCCCCAAATGGCAAGATTCAAGCCTTAAAGGAGATCTAACAAATGATAAG
TCAAGGTTCAAGCTTTTTACCTTGAATAAGCTGGAAATGAAGCAAAATCT
CTGGATCCACTTGCTTCTTCAAGAACCCCCAAGCTTCCACTTCTTCCTTC
AAGTTTCAAACAACTTTAAACACTCAAAAATGGCTCAAGAACTCAAAA
AGCTTTAGGGTTTCGAGTTAGGGCTTTTTGGAAGCGAGAGGGACGATGGG
40 GGCTGAAATGAGGCTAGAAAAAGTGTTTAAATAGGGGGCAAACCCTAAAT
ATTAGGGTTTCATCCAGGCAGCCCTACTCGTCGAGTCGGGCTCCCGACTC
GTCGAGTAGGTCACTTAAAACCCGCGTCCATAATCCAGTCTACTCGACGA
GTTGGGCCTCCAACCTCGTCGATTCCGAGTGCAAAACGTTCAATTACTTAA

ATTTAAATATGTACCAGGAACCGGGTGTACAGTTGAGACTTTATACCTC
CATAAGATAGATCTAGGTGCACATAGCCTGGATCCACAAGCTCCATGTCA
ACAAGCGACTCTTCAAGAAGTTCATTCTTCCTCCTTAAGCACCAAAAAAC
ACACAAAATCACCATGAAGCTCAAGAAATACTCAAATAGAGGATAGGGTT
5 TCGTTCGTAGGGTTAGAGAGGATGGAGGCTAGAGGAAATGAGGGATAGAG
GCGAGTTAAGGTCTTTAAATAGGGTCCAAGACCCTAAATTAGGGTTTTAA
TCTGGCCAGACGAACGCAGGGTGTTCCTCAAATGCATATGTGTCCAAATTC
TCGTGTGCGCCATGCGTACCTCCCTTGTACGCCATGTGTACCGGGTTTGG
TCCAAACCCTTCTAACTTCAAATGATCATAACTTGCACCCCTTATCTGTT
10 TTCGATGTTCTTTATATCCACGGAAAGGTAACAAGAAGCCCTATACTTCT
ATAAACTTTATTTAATCTGAAAACCAACCGAAATTAAATCCAAAATTCAT
AAAAGTCCCGAACCAACACATTTACCGATACCCTTGGGCTCCAAAACACA
AATTGAAAACCCGGATCATCCAACTACATCATCCACCTCCAAATGAGCC
CAAACCTCAATTATTCAAGGGTTCTAAGCCTGTTAATGCCCACTCCTCGAT
15 TACCACCCCGCAATGGGAAACGATTCAAACAGGGCGTTACATAATTTGT
TGTGGTTTTGTATTTTTTATTTCCGGTGAAGGTGAAAGATCCAATATTT
TTAATCTGTTGGCATTTCATCATTTGCAACTGTTTCTTGAAAAA
TACCTAAAATCAAATAACCATTTTCAAATCCAAAATTATAAGAGAGAAT
TGTAATGGACATGGAATCTTAAATCATTAAACACAGTTCAGTACACAAGT
20 TGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAAAGAG
ACATTACAAGAAGCCACTGACAGTATTTCTAATGTTGTATTCCCATCCTG
TCTCATGCACTCTTTTCATAACCTCCAGAACTTATATTGAACAGAGTTA
AAGGAGTGGAGGTGGTGTGTTGAGATAGAGAGTGAGAGTCCAACAAGTAGA
GAATTGGTAACAACCTCACCATAACCAACAACAGCCTGTTATATTTCCCAA
25 CCTCCAGCATTTGGATCTAAGGGGTATGGACAACATGATTTCGCGTGTGGA
AGTGCAGCAACTGGAATAAATTCCTCACTCTTCCAAAACAACAATCAGAA
TCCCCATTCCACAACCTCACAACCATAAATATTGATTTTTGCAGAAGCAT
TAAGTACTTGTTTTACCTCTCATGGCAGAACTTCTTTCCAACCTAAAGA
AAGTCAATATAAAATGGTGTGTTATGGTATTGAAGAAGTTGTTTCAAACAGA
30 GATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACACAACCAC
CATCTTGTTCCCTCATCTTGATTCTCTCACTCTAAGTTTCCTGGAGAATC
TGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAAATA
TCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCAATTTGAGGT
ATGCTTTGTTTCATATTCAATTATTTATTTAATTTCTTTTTTTATTTGCAA
35 TATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAATTAC
CTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAACTCTAGT
AAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGATGGG
TAATATAGGCAATTTAAGTTTTATTTCTGTAAAGCAGTATTTAGCAAGT
ACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGTCATT
40 GTACCCAGAATTTAGTTAAATGTAACATTTTAGATATCAGGGGACATCAG
GTGACAGATATTGTAGAATAGAACAAATATATAATATTACCCAAAACCTATT
TTTTCTAAGGTTTTTCTGTAAATATGTGCTTTCTTGATTTTCATTGAATT
TGCATTCCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAATCCC

GAAATTAATTAAAAAAAAAAAAAAAAAACAAAAGTAAATTTTTGATATGGAGA
GCACTGGTATCATTTAGTATATAAAAAAAGTAGATTTTGAATTAAGTTTC
TTATATAAAAGCTGTGTATATAGTTTAATTAGTTTACATCATTTTTCCA
TGTGGTGTTCAGTTGTCTGAAGCAGGTGGTGTTCCTTGGAGCTTATGCC
5 AATACGCTAGAGAGATAAGTATAGAATTCTGCAATGCATTGTCAAGTGTG
ATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGACAGT
CAGTTCTTGTAAATGGTCTGAAGGAGGTATTTGAAACTCAATTAAGGAGGA
GCAGCAACAAAAACAACGAGAAGAGTGGTTGTGATGAAGGAAATGGTGGA
ATTCCAAGAGTAAATAACAATGTTATTATGCTTTCTGGTCTGAAGATATT
10 GGAAATCAGCTTTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCAC
TTGAAAGCCTGAGACAGCTCGAAGAGTTAACGATAATGAATTGCTGGTCA
ATGAAAGTGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAAC
AACAAACAACGAAGGGGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
CTTCTTCTTCTTCTTCTTCTCCTCCTTCTTCTTCTAAGAAGGTTGTGGTC
15 TTTCTTGTCTAAAGTCCATTGTATTGGTCAATCTACCAGAGCTGGTAGG
ATTCTTCTTGGGGATGAATGAGTTCCGGTTCCTTCATTAGATGAACTTA
TCATCGAGAAATGCCCAAAATGATGGTGTTCACAGCTGGTGGGTCCACA
GCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACTATTGA
TCAAGAATCTGGCCTTAACCTTTCATCAGGTATATATGTTTCTTTAATTGG
20 CATCATCTAATTAAGAAAGATATCATTCCTGCCAAGTAAATTTACTTCAA
ACACATTACACTGGTTTTCAGTCTAAGTTTATGTTGTTCTAGGAAGGCCA
AAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTTCAAGTGGAAGGGTA
TTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGGCTTTTTAGTACCTG
GAATCGTGTGTGGGAGGAGCATTATTATTCTGATTTGCTTGTTCCTTTAT
25 CATTTTTTCTTAGCCTCTCGAACAGCTAGAAACCCTTTAATCTTTTGAT
TTTAAATGACAAAATTTTTCCCTGTTACTCTATTTGATTGTTGTTCTTCA
TGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCTTTTGATTGTTATT
TTCATAGCATGTTAGTCACTTGAATCAAGCTTTTTTCATTTTCAACCAGGG
CAAAAGGTCAAAAGTAACCTACTTTATGAGATCAAAAACAGCAACCCATC
30 GGATAACTTTTAGTTGGAGTTAATAGTTACAATTACCATTGTGATTAATA
ATTATAATATCCTGTATTAATTCATAAAAAATTGGTACAGCACATATATGA
CATTTCAAAGGTTTTTGTGTGACATATATGCCTCTGGCGTTTTCTTTA
TTGGACTTGCAGACCTCATTCCAAAGTTTATACGGTGACACCTTGGGCCC
TGCTACTTCAGAAGGGACAACCTTGGTCTTTTCATAACTTGATTGAATTAG
35 ATGTGAAATTTAATAAGGATGTTAAAAAGATTATTCCATCCAGTGAGTTG
CTGCAACTGCAAAAGCTGGAAAAGATAAATATAAACAGTTGTGTTGGGGT
AGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGGAGAAATGGAAATA
GTGGAATTGGTTTTGATGAATCGTCACAAACAACCTACCACTACTCTTGTC
AATCTTCAAACCTTAGAGAAATGAACTTATGGGGTCTAGATTGTCTGAG
40 GTATATATGGAAGAGCAATCAGTGGACAGCATTTGAGTTTCCAAAACATA
CAAGAGTTGAAATTAGTAATTGCAACAGTTTAGAACATGTATTTACTAGT
TCCATGGTTGGTAGTCTATCGCAACTCCAAGAGCTACATATAAGTCAGTG
CAAACTTATGGAGGAGGTGATTGTTAAGGATGCAGATGTTTCTGTAGAAG

AAGACAAAGAGAAAGAATCTGATGGCAAGATGAATAAGGAGATACTTGCG
TTACCTAGTCTAAAGTCCCTGAAATTAGAAAGCTTACCATCTCTTGAGGG
GTTTAGCTTGGGGAAGGAGGATTTTTTCATTCCCATTATTGGATACTTTAA
GAATTGAGGAATGCCCAGCAATAACCACCTTCACCAAGGGAAATTCCGCT
5 ACTCCACAATAAGAGAAATAGAAACAAGATTTGGCTCGGTTTATGCAGG
GGAAGACATCAAATCCTCTATTATAAAGATCAAACAACAGGTAAATCAGA
TCATTGTTGGTTTAATAATTCTTAAACTACATTTGAAAAGTTTCATGTAA
GTTTTTTATTATTGTCAAAAGCCGCAACCTATATTTTCAACTTTATATTT
ATGTACTTTATGCAGGATTTCAAAAAGCCCAGGACTCTATTTAATGTGA
10 AGTAAATACTAGAAGAGGTAAATTCTATTTACATGTCTCCTGATTGCCTA
TTAATTAATGGCCTTTCAGTTCATGGTTTTTGGATGTATTCTTCATGATG
ACGTGAATGTTTAAATACCCCACTAGTTAATTGTTAGGTTGAATGTTGAT
GACCAAAGGACTATATGTCTGGGAAGAATATTCAAGGAAAGAATTGTTTCAT
CATATGAAGGGCATTAATTAAGAAGAACATGGATGCTATGAAGATGTTG
15 GGAAAATATATGAATCAAATAACAAGCTACTCACTTATCTAAGTTTGTG
GTTGAGGATGTTGATTTTAATATTTCAAATTCATTGGTATCATTATATGG
GTTTATCAGTAGTGTTAATGGGATAATGAGCAACTTAACCTTAAATTATG
CTGTTGGTAAATGTTGGACTCAAGTATGGAAAATTAGGAATAACTTGTGA
AAAATATATGCAAAAAGTAGGATTGAGATTTTCAATGAAAAAATTATGAA
20 ACTATACTACTATAGTATATAAATAAATTCAACTTACTGTTGGGTATATT
GGAAGCACATATCATGAAAGTAACTAGAAGCAGAATTTGTTCCCATCTTC
ATCTACTTATAGTTTCCATTTCTTACTTGTAATAATCTGATTAACTTTA
GAGTTATTTCTATTTTTTACCAACCAAAATTTTCATATAAAGGCCACAAG
T (SEQ ID NO:107)

25

RG2J deduced polypeptide sequence (SEQ ID NO:108)

MSDPTGIVGAIINPIAQTALVPLTDHVGYMISCRKYVRDMQMKMTELNTRISAEEH
ISRNTRNHLQIPSIKDWLDQVEGIRANVANFPIDVISCCSLRIRHKLQKAFKITEQI
ESLTRQNSLIWTDEPVPLGRVGSMIASSTAASSDHHDVFPSSREQIFRKALEALEPVQ
30 KSHIALWGMGGVGKTTMMKKLKEVVEQKKTCNIIVQVVIGEKTNPAAIQQAVADY
LSIELKENTKEARADKLKRFEADGGKNKFLVILDDVWQFFDLEDIGLSPLPNKGV
NFKVLLTSRDSHVCTLMGAEANSILNIKVLKDEGKSLFRQFAKNAGDDDLDPFI
GIADSIASRCQGLPIAKTIALSLKGRSKSAWDVALSRLENHKIGSEEVVREVFKISYD
NLQDEVTKSIFLLCALFPEDFDIPIEELVRYGWGLKLFIEAKTIREARNRLNNCTERL
35 RETNLLFGSHDFGCVKMHDVVRDFVLHMFSEVKHASIVNHGNMSEWPEKNDTSN
SCKRISLTCKGMSKFPKDINYPNLLILKLMHGDKSLCFPENFYGKMEKVQVISYDKL
MYPLLPSSLECSTNVRVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIG
NLKKLRLLDLTNCKGLRIDNGVLKNLVKLEELYMGVNRPYQAVSLTDENCNEM
VEGSKKLLALEYELFKYNAQVKNISFENLKRKISVGC SLHGSFSKSRHSYENTLKL
40 AIDKGELLESRMNGLFEKTEVLCLSVGDMYHLSDVKVKSSSFYNLRVLVSECAEL
KHLFTLGVANTLSKLEHLKVYKCDNMEELIHTGGSEGDTITFPKLKLLYLHGLPNL
LGLCLNVNAIELPKLVQMKLYSIPGFTSIYPRNKLEASSLLKEEVVPEELIVEKCGSI
EELFNIDLDCASVIGEEDNNSLRNINVENSMKLEREVWRIKGADNSRPLFRGFQVVE

KIIITRCKRFTNVFTPTTNFDLGALLEISVDCRGNDESDQSNQEQEQIEILSEKETLQE
ATDSISNVVFPSCLMHSFHNLQKLILNRVKGVVVFEIESESPTSRELVTTHHNQQQP
VIFPNLQHLLDRGMDNMIRVWKCSNWNKFFTLPKQQSESPFHNLTINIDFCRSIKY
LFSPLMAELLSNLKKVNIKWCYGIEEVVSNRDDDEDEEMTTFTSTHTTTILFPHLDSL
5 TLSFLENLKCIGGGGAKDEGSNEISFNNTTATTA VLDQFELSEAGGVSWSLCQYAR
EISIEFCNALSSVIPCYAAGQMQLQVLTVSSCNGLKEVFETQLRRSSNKNNEKSGC
DEGNNGGIPRVNNNVIMLSGLKILEISFCGGLEHIFTFSALSLRQLEELTIMNCWSMK
VIVKKEEDEYGEQQT TTTTKGTSSSSSSSSSSSSSSSSSSSSSKKV VVFPCLKSIVLVNLP
ELVGFFLGMNEFRLPSLDELIEKCPKMMVFTAGGSTAPQLKYIHTRLGKHTIDQES
10 GLNFHQDIY MPLAFSLDLQTSFQSLYGD TLGPATSEGTTWSFHNLIELDVKFNKD
VKKIIPSELLQLQKLEKININSCVGVVEVFETALEAAGRNGNSGIGFDESSQTTTTTL
VNLPNLREMNLWGLDCLRYIWKSNQWTA FEFPKLTRVEISNCNSLEHVFTSSMVGS
LSQLQELHISQCKLMEEVIVKDA DVSV EEDKEKESDGKMNKEILALPSLKS LKLESL
PSLEGFSLGKEDFSFPLD TLRIECPAITTF TKGNSATPQLREIETRFGSVYAGEDIKS
15 SIIKIKQQDFKKAQDSI.CEVNTR

RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110)

TGGGATTCCATATATAAAAACATATATTTTTATAAAGTGGGATTCCATTG
TTTATATAGATTTTTATTACCAATAGACAATAGATTAAAAAAGATATA
20 AAAACATGTCGGCTTTTGACTAAAAATATAGATTTTTATGAATAGAATAT
TCAATTTGCTTAACTCGTTTAAAAAAAATGAAAAAGATGTCGATATAAAA
TCTCATATGGGCCTTCTTTACCATTCAAATAGTAAAATAGTAAAAGATAC
TTGTTTGGGGCATGAACTGACCATAGTCAAACCCATACAAAATCAAACGA
ATCCACATGGATGATGACGATGGGGTCGCAGTAAATGTGTTTTGGTCCT
25 TTTTTTTCGAGAGAACAGAAGCTTCTGCTCTTCATCTTCTTTAGATTTTG
GGGATTTTCTGGTTTCAGGGGTTTGTGAGTGGAACATAAATTGAAGCAA
AAAGTATGGTATAATTGGTTGCTAGTGAAATTGATGCTTTCTATTACTAT
CATCTTTAAAATTGTCAAAACATTATGTATTAAATTATGAGATCGAAAGT
GGTCTATGGGCCAAAGGTAATACAAGCTTACTCAATGAAATGAATCTAGG
30 ATGCATCATGCATGTATTGGTTAGATTAAAGATTTTCATCAAATTTCTCT
TATCAAATTGTTGTATAACCATGTTATGTAGGTGCTACCACAAGCCATAAC
ATCGAGCAATGGAGTGTATTACTGGCATCTTTAGCAACCCGTTTGCTCAG
TGTCTCATCGCTCCTGTGAAAGAACACCTTTGCCTTCTGATTTTCTATAC
ACAATATGTAGGGGATATGCTTACTGCAATGACGGAGTTGAATGCTGCAA
35 AAGACATTGTTGAAGAGCGGAAGAATCAAAACGTAGAAAAATGTTTTGAG
GTTCCAAACCATGTCAACCGTTGGTTGGAAGATGTTCAAACAATCAACAG
AAAAGTGGAACGTGTTCTTAACGATAATTGCAATTGGTTCAATCTATGTA
ATAGGTACATGCTCGCAGTGAAAGCCTTGGAGATAACTCAGGAGATCGAT
CATGCCATGAAACAACCTCTCTCGGATAGAATGGACTGATGATTCAGTTCC
40 TTTGGGAAGAAATGATTCCACAAAGGCATCCACCTCTACACCATCAAGTG
ATTACAATGACTTCGAGTCAAGAGAACACACTTTTAGGAAAGCACTTGAA
GCACTTGATCCAACCACACATCCACATGGTAGCCTTATGGGGGATGGG

TGGAGTTGGGAAGACCACGATGATGAAGAGGCTGAAAAATATTATTAAAG
AAAAGAGGACGTTTCATTATATTGTTTTGGTGGTTATAAAGGAAAATATG
GATCTCATTTCCATCCAGGATGCTGTAGCAGATTATCTGGATATGAAGCT
AACAGAAAGCAATGAATCAGAAAGAGCCGATAAACTTCGTGAAGGGTTTC
5 AGGCCAAATCAGATGGAGGTAAGAATAGGTTCCCTCATAATACTGGATGAT
GTATGGCAATCTGTTAATATGGAAGATATTGGTTTAAGTCCTTTTCCGAA
TCAAGGTGTCGACTTCAAGGTCTTGTTGACCTCGGAAAACAAAGATGTTT
GTGCAAAAATGGGAGTTGAAGCTAATTTAATTTTCGACGTGAAATCTTA
ACAGAAGAAGAAGCACAAAGTTTGTTTTATCAATTTGTAAAAGTTTCTGA
10 TACCCACCTTGATAAGATTGGAAAAGCTATTGTAAGAACTGTGGTGGTC
TACCCATTGCCATCAAAACCATAGCCAATACTCTTAAAAATAGAAACAAG
GATGTATGGAAGGATGCACTTTCTCGTATAGAGCATCATGACATTGAGAC
AATTGCACATGTTGTTTTTCAAATGAGCTACGACAATCTCCAAAACGAAG
AAGCTCAATCCATTTTTTTTGCTTTGTGGATTGTTTCCTGAAGACTTTGAT
15 ATTCCTACTGAGGAATTGGTGAGGTATGGATGGGGATTGAGAGTATTTAA
TGGAGTGTATACTATAGGAGAAGCAAGACACAGGTTGAACGCCTACATCG
AGCTGCTCAAGGATTCTAATTTATTGATTGAAAGTGATGATGTTCACTGC
ATCAAGATGCATGATTTAGTTCTGTGCTTTTGTTTGGATACGTTTAATAG
ATTCAAGCATTCTTTGATTGTTAACCATGGTAATGGTGGTATGTTAGGGT
20 GGCCTGAAAATGATATGAGTGCCTCATCTTGCAAAAGAATTTCATTAATA
TGC.AAGGGCATGTCCGATTTTCTTAGAGACGTAAAGTTTCCAAATCTCTT
GATTTTGAAACTTATGCATGCAGATAAGTCTTTGAAGTTTCTCAAGACT
TTTATGGAGAAATGAAGAAGCTTCAGGTTATATCATACGATCACATGAAG
TATCCCTTGCTTCCAACATCACCTCAATGCTCCACCAACCTTCGTGTGCT
25 TCATCTTCATCAATGCTCATTGATGTTTGATTGCTCTTCTATTGGAAATC
TGTTGAATCTGGAAGTGCTCAGCTTTGCTAATTCTGGTATTGAGTGGTTG
CCTTCCACAATCGGAAATTTGAAGGAGCTAAGGGTACTAGATTTGACAAA
TTGTGATGGTCTTCGTATAGATAATGGTGTCTTAAAGAAATTGGTGAAAC
TTG.AAGAGCTTTATATGAGAGTTGGTGGTCGATATCAAAAGGCCATTAGC
30 TTC.ACTGATGAAAACCTGCAATGAAATGGCAGAGCGTTCAAAAAATCTTTC
TGC.ATTAGAATTTGAGTTCCTTCAAAAACAATGCTCAACCAAAGAATATGT
CATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGATGTTATTTAAG
GGAGATTTTCGGTAAGATCTTTCCTCTTTTGAACACGTTGCGGTTGGT
CACCAACAGAACTGAAGTTCTTGAATCTAGGCTTAATGAGTTGTTGAGA
35 AAACAGATGTTCTTTATTTAAGTGTGGGAGATATGAATGATCTTGAAGAT
GTTGAGGTAAAGTTGGCACATCTTCCTAAATCCTCTTCCTTCCACAATTT
AAGAGTCCTTATCATTCTGAGTGTATAGAGTTGAGATACCTTTTCACAC
TTG.ATGTTGCAAACACTTTGTCAAAGCTTGAGCATCTTCAAGTTTACGAA
TGCGATAATATGGAAGAAATCATACATACAGAGGGTAGAGGAGAAGTGAC
40 AATTACATTCCCAAAGCTGAAGTTTTTATCATTGTGTGGGCTACCAAATC
TGTTGGGTTTGTGTGGTAATGTGCACATAATTAATCTACCACAACCTACA
GAGTTGAAACTTAATGGCATTCCAGGTTTCACAAGCATATATCCTGAAAA
AGATGTTGAAACATCTAGTTTGTGAATAAAGAGGTAATGTGTTTTATG

TTAATACAATACAATCTTTTCAATTAACCGTTTCAAAAATATATTGTATGA
TTTATTTTTGTTTGGATGGGGTTATTAATGGGTGATTATTTCTCAGGTTG
TAATTCCTAATTTGGAGAACTTGATATTAGTTATATGAAGGATTTGAAA
GAGATATGGCCTTGTGAATTAGGGATGAGTCAGGAAGTTGATGTTTCTAC
5 GTTGAGAGTGATTAAAGTAAGCAGTTGTGATAATCTTGTGAATCTATTCC
CGTGCAATCCTATGCCATTGATACATCACCTTGAAGAGCTTCAAGTGATA
TTTTGTGGTTCATTGAAGTGTTATTCAACATTGAGTTGGATTCTATTGG
TCAAATTGGAGAAGGCATCAACAATAGCAGCTTGAGAATCATCCAATTGC
AGAACTTAGGGAAGCTAAGTGAGGTGTGGAGGATAAAAGGTGCGGATAAC
10 TCTAGTCTTCTCATCAGTGGCTTTCAAGGTGTTGAAAGCATTATCGTTAA
CAAATGCAAGATGTTTAGAAATGTATTCACACCTACCACCACCAATTTTG
ATCTGGGGGCACCTTATGGAGATTCGGATACAAGATTGTGGAGAAAAGAGG
AGAAACAACGAATTGGTAGAGAGTAGCCAAGAGCAAGAGCAGGTATGGCT
TTCAATTTCACTTTCTTACTTAATGAAGGATTAAGCTCCTGCTTTTTTGAA
15 TAAAAAGTGGATGAATGACTAAATTCGGGAATGCCACCCGGAAAGTTATC
AACCATTTAGCTACACCATTTTTTTGAACTAATGTTGCAATAAATGCATAA
TATAATTAAAAAATGGTCATTGATAAATGTAAACCAACCTTTTTTATTTA
TTAAAATGTCTACAATAAATGATTTTCTTTATTATATATCATTTTATAAC
AAT.AAGCTTAAAGATGTTTAAATAGCCAATGTCAGTTATAGATCGTAACT
20 AATTTTTTATTAAGTATGTTTATAGTAAAGATATCACTCATTATTATTTTAA
TAGAAAAAAGACAAGATTGGCTAATCCTCATAAGAATTTGGAAGATTTAA
GCAAAATATAGAGCTTTTCCAAACATAGCCAATAGTTTCTTTTGCAGGTC
CCATCTACGAAATTATCAATAGATTTGCGATTTTTTTTTTGGCACCCGGGA
AATTTCCATTAATTAATAAAAAAAGTTCAAGCCATTTTGAGTTGGCACCTG
25 CAAAATGGTAGTTTGCACCTGCGGAAATCACCTTTCACCATTTTCGCATCT
ATG.ACTTGTGAAAATGTTAATTTGTGAAATGGTCATGTGCACCTCATGAG
AAATACGAAATGGTCAGTAATATGACTTTTTTATATAAATATGATGGTGG
CATATATTTATAGGAAAATATAGCTGCACGATATTAATTAATAGTGAAAT
TAGTTAACTGTATACGATAAGTATACAAAATTTATATGTATGAAGTATAC
30 TCAATTTAGGACGACTCGGGCAATGAAATCATCATTTAATAGGAGCAATG
AAATCATTTTCGAAAAATGTTTACAAATGAATAAAATATTAAATTAACT
TAAACATTTTGTTAGTAGTTTGAAATTTACAAACTGAAATTTGTTGTAT
TTATTAACATTTATAAATGTTGTACTATGATTTTTTCCTTGTTTGCAAAT
ATTCCTTAAAAATCCACCTAAAATCAAAATAATTAATCTTTTTCAAGTTG
35 AAAAATGAAAATCGTATGATATAACCGTGTATGGATGTGGAATTATATAT
CAGTTACTAATTACATTTTTTTGTTGGGATATATGTGCGCAGATTGATATT
GCAATCCCATTCACTCTCACACACTCTTCCAAAACCTCCGTAAACTTGC
TTTGGAAAAGTATGAAGGAGTGGAGGTGGTGTGTTGAGATAGAGAGTCCAA
CAAGTAGAGAATTGATAACAATTCACCATAATCAACAACCACTACTTCCC
40 AACCTTGAGTTATTGGATATAAGTTTTATGGACAGCATGAGTCATGTATG
GAAGTGCAACTGGAATAAATCTTCATTCTTCAAAAACAACAGTCAGAAT
CCCCATTCTGTAATCTCACAACCATACATATTCAATATTGCCAAAGCATT
AAGTACTTGTTTTCAACTCTCATGGCAAACTTCTTTCCAACCTAAAGAA

GGTCGAGGTAAGAGAGTGTTCATGGTATTGAAGAAGTTGTTTCGAACAGAG
ATGATGAAGATGAGGAAAAGACTACATTTACATCTACATCTTCTGAAAAA
AGCACTAATTTGTTCCCTCGTCTTGAATCTCTCGCTCTTTATCAACTTCC
AAATCTCAAGTGTATTGGTGGTGGTGGTCTGCCAACAGTGGGAACAATG
5 AAATATCTCTTGATAATTCCACTACTACTACTTCTTTTGTGATCAATCT
AAGGTATGTTTTTTTTTTNGTTNCCCTT (SEQ ID NO:109)
Sequence gap
CCTCCCTAATAATACATGTTATGCACACTATACTAACATATTAGACACGT
AAAGGATAAATGCTATGCCTCATATAATACGTTATATTTATAATCTTTAA
10 ACAATCAAATTTATTAAACAAATAACTAAGTGTGAGCAAAGGCAGGTACC
CGACTAAATTGCCCAAACCAGTCTGGTGGTTCGTGGAATGTTGGGCCAG
GTCGTTAAACGTCTACACACCGGTTCTTTAAATCACAGATCCGCTTCTC
ATACTGTGAACCCGGTTTTAATTTTAAAAGAAAATTCATTATAAAGTAA
ATGACTTAAACCATTACAAACAACAAAAATTTACCATTACAATGTTGGAC
15 TATCATTATTTGCAACATAAACTGAAAATACACATATTTCTTCTGATA
TCAGCATGAGTGGCTGGTTGGCTAACCCAAAAATCCATGCATTGTAGATG
TGTGTTACAACACATAGTATCAATGAAAGGCATATTTTAGGCTAGAATT
TAACAATCTGTAATAATATCCCTAAAATAATATCATCATCAACCAACT
AATATAAAACCATTGGGTTTCGTCAATTTTAGGTACAAAACATAGATTTTC
20 TAAGCTTGTTGTATTTAAACATATGCTTTCTAACTTAATTGATTTTGCA
TTCCAAAATTTTAGGTTGTAAAGTGGTATGTCATTTGTTGTCTTTTCAAC
ATTAAATTGTACAAAAACCAAACTACATAATTGATGTAGATATCATAACA
ATTGTGTTATTTAGTATATAAAAACTAAATTTGAATTGAATTTCTTATA
CAAAAGTTGTGTCTATGTATACATGTTTATGTAGGTAATAGACAATTAGT
25 CTCTGTTAAGTATATGGAGTTTAATTTTAGACTAATTTTCATGTGTTG
CAGTTTTATCAGGCAGGTGGCGTTTTTTGGACGTTATGCCAATACTCCAG
AGAGATAAATATAAGGGAGTGTTATGCATTGTCAAGTGTAATTCCATGTT
ATGCAGCAGGACAGATGCAAAATGTTCAAGTGCTGAATATATACAGGTGC
AACTCAATGAAGGAGTTATTTGAACTCAAGGGATGAACAACAACAATGG
30 TGACAGTGGTTGTGATGAAGGAAATGGTTGTATACCAGCAATTCCAAGAC
TAAATAACGTTATTATGCTACCCAATCTAAAGATATTGAAGATTGAAGAT
TGTGGTCATCTGGAACATGTATTCACATTCTCTGCACTTGGAAGCCTGAG
ACAGCTCGAAGAGTTAACGATAGAGAAATGCAAGGCAATGAAAGTGATAG
TGAAGGAAGAAGATGAATATGGAGAGCAAACAACAAGGCATCTTCGAAG
35 GAGGTTGTGGTCTTTCCCTCGTCTCAAGTCCATTGAACTGGAAAATCTACA
AGAGCTCATGGGTTTCTACTTAGGGAAGAATGAGATTCAGTGGCCTTCAT
TGGATAAGGTTATGATCAAGAATTGCCCAGAAATGATGGTGTGTTGCACCT
GGTGAGTCCACAGTTCCCAAGCGCAAGTATATAAATACAAGCTTTGGCAT
ATATGGGATGGAGGAGGTACTTGAACTCAAGGGATGAACAACAATAATG
40 ATGACAATTGTTGTGATGATGGAAATGGTGGAAATCCAAGACTAAATAAC
GTTATTATGTTTCCAAATATAAAGATATTGCAAATCAGCAATTGTGGCAG
TTTGGAACATATATTCACATTCTCTGCACTTGAAAGCCTGATGCAGCTCA
AAGAGTTAACAATAGCGGATTGCAAGGCAATGAAAGTGATTGTGAAGGAG

GAATATGATGTAGAGCAAACAAGGGTATTGAAGGCTGTGGTATTTTCTTG
TCTAAAGTCCATTACACTATGCCATCTACCAGAGTTGGTGGGTTTCTTCT
TGGGGAAGAATGAGTTCTGGTGGCCTTCATTGGATAAGGTTACCATCATT
GATTGCCCACAAATGATGGGGTTCACACCTGGTGGGTCAACAACCTCCCA
5 CCTCAAGTACATACTCAAGCTTAGGCAAACATACTCTTGAATGTGGCC
TTAATTTCAAGTCACAACACTGCATATCATCAGGTATAATTATTATTCT
TTNACACCATCTAATTATGGAATCATGACGCTAATTACAGTATTAAACAC
(SEQ ID NO:110)

10 **RG2K deduced polypeptide sequence (SEQ ID NO:111)**

MECITGIFSNPFAQCLIAPVKEHLCLLIFYTQYVGDMLTAMTELNAKDIVEERK
NQNVKEKCFEVPNHVNRWLEDVQTINRKVERVLNDNCNWFNLCNRYMLAVKAL
EITQEIDHAMKQLSRIEWTDDSVPLGRNDSTKASTSTPSSDYNDFESREHTFRKAL
EALGSNHTSHMVALWGMGGVGKTTMMKRLKNIKEKRTFHYIVLVVIKENMDL
15 ISIQDAVADYLDMKLTESNESERADKLREGFQAKSDGGKNRFLIILDDVWQSVN
MEDIGLSPFPNQGVDFKVLITSENKDVCAKMGVEANLIFDVKFLTEEEAQSLFY
QFVKVSDTHLDKIGKAIVRNCGGLPIAKTIAN TLKRNKDVWKDALSRIEHHD
IETLAHVVFQMSYDNLQNEEAQSIFLLCGLFPEDFDIPTEELVRYGWGLRVFNGV
YTIGEARHRLNAYIELLKDSNLLIESDDVHCIMHDLVRAFVLDTFNRFKHSLIV
20 NHGNGGMLGWPENDMSASSCKRISLICKGMSDFPRDVKFPNLLILKLMHADKS
LKFPQDFYGEEMKKLQVISYDHMKYPLLPTSPQCSTNLRVLHLHQCSLMFDCSSI
GNLLNLEVLFSFANGIEWLPSTIGNLKELRVLDLTNCDGLRIDNGVLKKLVKLEELY
MRVGGRYQKAISFTDENCNEMAERSKNLSALEFEFFKNNAPKNMSFENLERFKIS
VGCFYKGDGFKIFHSFENTLRLVTNRTEVLESRLNELFEKTDVLYLSVGDMDNDLED
25 VEVKLAHLPKSSSFHNLRLVLISECIELRYLFTLDVANTLSKLEHLQVYECDNMEEII
HTEGRGEVTITFPKLKFLSLCGLPNLLGLCGNVHIINLPQLTELKLNIGIPGFTSIYPEK
DVETSSLLNKEVVIPNLEKLDISYMKDLKEIWPCELGMSQEV DVSTLRVIKVSSCDN
LVNLFPCNPMPLIHHLEELQVIFCGSIEVLFNIELDSIGQIGEGINSSLRRIQLQNLGK
LSEVWRIKGADNSSLLISGFQGVESIVNKCKMFRNVFTPTTTNFDLGALMEIRIQDC
30 GEKRRNNELVESSQEQEQ

RG2L polynucleotide sequence (SEQ ID NO:112)

GGAAGACACAATGATGCAAAGACTGAAGAAGGTTGCCAAAGAAAATAGAA
TGTTTCAGTTACATGGTTCGAGGCAGTTATAGGGGAAAAGACAGACCCAATT
35 GCTATTCAACAAGCTGTAGCCGATTACCTTCGTATACAGTTCAAAGAAAG
CACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCCCCACT
CTGNAGACGGTAAGAATAAGTTCCTCGTAATATTTGATGACGTCTGGCAG
TCCGTTGATCTGGAAGATATTGGNTTAAGTCCTTTTCCAAATCAAGGTGT
CGACTTCAAGGTCTTGTTGACTTCACGAGACGAACACGTTTGCACAATGA
40 TGGGGGTTGAAGCTAATTCAGTTATTAATGTGGGACTTCTAACTGAAGTA
GAAGCACAAAGTCTGTTCCAGCAATTTGTAGAAACTTTTGAGCCCCGAGCT
CTGTAAGATAGGAGAAGTTATCGTAAGAAAGTGTGCGGTCTACCTATTG

CCATCAAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGG
AAGGATGCACTTTCACGTATAGAGCACTATGACATTCGTAGTGTTCGCC
TAAAGTCTTTGAAACAAGCTATCACAATCTCCAAGACAGGGAGACTAAAT
CCGTGTTTTTGATGTGTGGTTTGTTCCTGAAGACTTCAATATTCCTACC
5 GAGGAGTTGATGAGGTATGGATGGGGCTTAAAGCTATTTGACAGAGTTTA
TACAATTAGAGAAGCAAGAACCAGGCTCAACACCTGCATTGAGCGACTTG
TGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATG
CATGATCTGGTGCCTGCTTTTGTTCCTGAAGTTCGAGCA
TGCTTCAATTGTCAACCATGGTAATATGCATGGGTGGACTAAAAATGATA
10 TGAACGACTCTTGCAAAACAGTTTCTTTAACATGCGAGAGTGTGTCTGAG
TTTCCAGGAGACCTCAAGTTTCCAAACCTAAAGCTTTTGAACTTATGCA
TGGAGATAAGATGCTAAGGTTTTCTCAAGACTTTTATGAAGGAATGGAAA
AGCTCCAGGTAATATCATACCATAAAATGAAGTATCCATTGCTTCCCTCG
TCACCTCAATGCTCCACCAACCTTCGAGTGCTTCATCTTCATCGGTGTTT
15 ATTACGGATGCTTGATTGCTCTTGATCGGAAATTTGACGAATCTGGAAG
TGTTGAGCTTCGCTAATTCTGGCATTGAACGGATACCTTCAGCAATCGGA
AATTTGAAGAAGCTTAGGCAACTTGATCTGAGAGGTCGTTATGGTCTTTG
TATAGAACAGGGTGTCTTGAAAAATTTGGTCGAACTTGAAGAACTTTATA
TTGGAAATGCATCTGCGTTTAGAGATTATAACTGCAATGAGATGGCAG
20

RG2L deduced polypeptide sequence (SEQ ID NO:113)

EDTMMQRLKKVAKENRMFSYMVEAVIGEKTDPPIAQQAVADYLRIQFKESTKPAR
ADKLREWFKAHS?DGKNKFLVIFDDVWQSVLDLEDIGLSPFPNQGVDFKVLLTSRDE
HVC TMMGVEANSVINVGLLTEVEAQSLFQQFVETFEPELCKIGEIVVRKCCGLPIAI
25 KTMACTLRNKRKDAWKDALSRIEHYDIRSVAPKVFETSYHNLQDRETKSVFLMCG
LFPEDFNIPTEELMRYGWGLKLFDRVYTIREARLNTCIERLVQTNLLIESDDVGC
VKMHDLVRAFVLGMYSEVEHASIVNHGNMHGWTKNMNDSCKTVSLTCESVSEF
PGDLKFPNLKLLKLMHGDKMLRFSQDFYEGMEKLQVISYHKMKYPLLPSPQCST
NLRVLHLHRCSLRMLDCSCIGNLTNLEVLSFANS GIERIPSAIGNLKKLRQLDLRGR
30 YGLCIEQGV LKNLVELEELYIGNASAFRDYNCNEMA

RG2M polynucleotide sequence (SEQ ID NO:114)

GGGGAAGACACAATAGATGCAAAGGCTGAAGAAGTTGCCAAAGAAAAGAG
AATGTTCAAGTTATATCATTGAGGCGGTTATAGGGGAAAAGACAGACCCCA
35 TTTCCATTCAAGGAAGCTATATCATATTACCTTGGTGTAGAGCTCAATGCA
AATACTAAGTCAGTAAGAGCTGATATGCTTCGTCAAGGGTTCAAGGCCAA
ATCTGATGTAGGTAAGGATAAATTCTTAATAATACTCGACGATGTATGGC
AGTCTGTTGATTTGGAAGATATTGGATTAAAGTCCATTTCCAAATCAAGGT
GTTAACTTCAAGGTCCTGTTAACATCACGAGACCGACATATTTGCACTGT
40 GATGGGGGTTGAAGGTCATTGATTTTAAATGTGGGACTTCTCACAGAAG
CAGAATCAAAAAGATTGTTCTGGCAGTTTGTAGAAGGTTCTGATCCTGAG
CTCCATAAGATAGGAGAAGATATTGTAAGTAAGTGTGTGGTCTACCCAT

TGCCATTAAAACCATGGCATGTACACTTAGAGATAAAAGTACGGATGCAT
GGAAGGATGCACTGTCTCGTTTAGAGCATCATGACATTGAAAATGTTGCC
TCTAAAGTTTTTAGAGCGAGCTATGACCATCTCCAAGACGAGGAGACTAA
ATCCACTTTTTTCTATGTGGATTGTTTCCAGAAGATTCCAATATTCCTA
5 TGGAGGAGTTGGTGAGGTATGGGTGGGGATTGAAATTATTTAAAAAAGTG
TATACCATAAGAGAAGCAAGAACTAGGCTCAACACTTGCAATTGAGCGGCT
CATCTATACCAATTTGTTGATAAAAGTTGATGATGTTTCAGTGCATCAAGA
TGCATGATCTCATCCGTTCTTTTGTGTTTGGATATGTTTTCTAAAGTTGAG
CATGCTTCGATTGTCAACCATGGTAATACGCTAGAGTGGCCTGCAGATNA
10 TNTGCACGACTCTTGTAAGGGCTTTCATTAAACATGCAAGGGTANATGTG
AGTTTTGTGGAGACCTNAANTTTCCAACCCTAATGATTTTAAACTTATG
CATGGAGATAAATCGCTAAGGTTT

RG2M deduced polypeptide sequence (SEQ ID NO:115)

15 GEDTIDAKAEVAKERMFYSYIEAVIGEKTDPISIQEAISSYYLGVELNANTKSVRAD
MLRQGFKAASDVGKDKFLILDDVWQSVLEDIGLSPFPNQGVNFKVLLTSRDRHI
CTVMGVEGHSIFNVGLLTEAESKRLFWQFVEGSDPELHKIGEDIVSKCCGLPIAKT
MACTLRDKSTDWKDALSRLHHDIENVASKVFRASYDHLQDEETKSTFFLCGLFP
EDSNIPMEELVRYGWGLKLFKKVYTIREARLNTCIERLIYTNLLIKVDDVQCIKM
20 HDLIRSFVLDMFSKVEHASIVNHGNTLEWPAD??HDSCKGLSLTCKG?CEFCGDL?F
PTLMILKLMHGDKSLRF

RG2N polynucleotide sequence (SEQ ID NO:116)

AGGTAAATCCATAACCCCTAAATGTTGGTACGCTCATATATCAAATTGCG
25 TGTTTTGTTGAATGAAAAAGCATGCTCAAAAAACCAGTGTAAGGCACGG
TATATGACATATTTATAGTTACTGATAACAAATTATGATAATTTGGGTT
TACRTAAGTTAGGATTCGTACTTCAACCAAATGTAATAGTTTTTGTGAGT
CTATCTATGTATTTGGGGAATCACATTAGCAACGGGATTGTACTAGTAAT
TCGAAAAAGTCTTTTAAATAATTTTCTGTTTATAATTTATGAATAGTTT
30 TAGCGACATCTAATATTAATAGAAATGTATCTGATATTGAATTAATGTCC
TTAATGTGAACATAGACCTTTTCCATTTACTAATGCCTAATTATTAGTTT
CTAATCAATAAATTTTAATTTCTGTTTTATGCTTCTAAGACAATAAAAAAT
CCATGATTTACCTTTAAATATTAACAAAAATGACCATAAATAAATAAAAA
ATTAGGATACCAAACCCCCCGCCATGCCCAATGTCTAAATATTCTTGAT
35 GCTTTTGCTTTTCCCTCTTTTCCCTTGTTAGTCTATTATTCTGGAGAGTTT
GAGAGAGTTTCATACAAGAAAATTTCAAGAAGAAAGCAAAGGTCCAGGTA
TTCTCTTTTCTTAATTATGTATTAACCTTACAAGCATTTTTTACACGATCC
ATGGTTTTTTGTGTATGTTTTTCAAATTGAACTAGATTGGGACTTTTGC
CCTTGATGATTCATAAGATATTGCATGGAGTTGAGATTGTGTAAGAAAAG
40 TGGTGAATAGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATATG
ATGATGAGATAGAGATATGTTAAACTGGCTAGAAAATTGTTTTAATTTG
AAATTTAGGTKGTTGAATTTGAAAGATACCAAGCTAATAACTAATTAGTT

ATGCTAAWTAGTTATAAAGAACAACAAACTCTTAGTTTTTTTTTTCATGA
TTTTCAACCTCTTTGTACCAAATAAATTATAGCAAAATTGAATATCATT
CTCTGCAATCAATCTTAACTTTTGTTATTATCATCATGTCTAAAATTGCC
ACAAGTTTATTTTCAAAGTCATATTGGATTATGAAAGGACTATTTTTACC
5 AATTACATCTTTACTTTATGGGCCAAAGCTAATACAATCCGACTAAACTA
AAGGAATATGGGATGCATATAGTTTGCTTCCCGATTATAGATTTCTATCT
AATTTGTCTATTGTACTAATTTAGGTGCCACCACAAGTAAATTTGTTAAA
TGGATATCGTTAATGCCATTCTTAAACCAGTTGTCGAGACTCTCATGGTA
CCCGTTAAGAAACACATAGGGTACCTCATTTCCCTGCAGGCAATATATGAG
10 GGAAATGGGTATCAAAATGAGGGGATTGAATGCTACTAGACTTGGTGTCG
AAGAGCATGTGAACCGGAACATAAGCAACCAGCTTGAGGTTCCAGCCCAA
GGCAGGGGTTGGTATGAAGAAGTAGGAAAGATCAATGCAAAAGTGGA AAA
TTTTCTAGCGATGTTGGCAGTTGTTTCAATCTTAAGGTTAGACACGGGG
TCGGAAAGAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTCTAGAGA
15 GAACACTCTATCATCATCTGGAATGATCATTCCATTCTTCTAGGAAGAAT
TGATTCCACGAAAGCATCCACCTCAATACCATCAACCGATCATCATGATG
AGTTCCAGTCAAGAGAGCAAACCTTTCACAGAAGCACTAAACGCACTCGAT
CCTAACCACAAATCCCACATGATAGCCTTATGGGGAATGGGCGGAGTGGG
GAAGACGACAATGATGCATCGGCTGAAAAAGGTTGTGAAAGAAAAGAAAA
20 TGTTTAATTTTATTGTTGAGGCGGTTGTAGGGGAAAAAACAGACCCCAT
GCTATTCAATCAGCTGTGGCAGATTACCTAGGTATAGAGCTCAATGAAAA
AACTAAACCAGCAAGAACTGAGAAGCTTCGTAAATGGTTTGTGGACAATT
CTGCTGGTAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTTGT
GATCTGAATGATATTGGTTTAAGTCCTTTACCAAATCAAGGTGTGCGACT
25 CAAGGTGTTGTTGACATCACGAGACAAAGATGTTTGCAGTGAAGTGGGAG
CTGAAGTTAATTCAACTTTTAATGTGAAAATGTTAATAGAAACAGAAGCA
CAAAGTTTATTCCACCAATTTGTAGAAATTTTCGGATGATGTTGATCGTGA
GCTCCATAATATAGGAGTGAATATTGTAAGGAAGTGTGGCGGTCTACCCA
TTGTCATCAAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCA
30 TGGAAGAATGCACTTCTTCGTTTAGTGAACATAACATTGAAAATATAGT
GAATGGAGTTTTTAAAATGAGTTACGACAATCTCCAAGATGAGGAGACTA
AATCCACCTTTTTGCTTTGTGGAATGTTTCCCGAAGACTTTAATATTCCT
ACCGAGGAGTTGGTGAGGTATGGATGGGGGTTGAAATTATTTAAAAAAGT
GTATACTATAGGAGAAGCAAGAATCAGGCTCAACACATGCATTGAGCGGC
35 TCATTATACAAATTTGTTGATTGAAGTTGATGATGTTAGGTGCATCAAG
ATGCATGATCTTGTCCGTGCTTTTGTGTTGGATATGTATTCTAAAGTCGA
GCATGCTTCCATTGTCAACCATGGTAATACACTAGAGTGGCATGTGGATA
ATATGCACAACTCTTGTAAGAACTTTCATTAAACATGCAAGGGTATGTCT
AAGTTTCCTACAGACCTCAAGTTTCCAAACCTCTCGATTTTGAACTTAT
40 GCATGAAGATATATCATTGAGGTTTCCCAAAAACCTTTTATGAAGAAATGG
AGAAGCTTGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCC
TCATCACCGCAATGCTCCGTCAACCTTTGCGTGTTTCATCTCCATAAATG
CTCGTTAGTGATGTTTGACTGCTCTTGATTGGAAATCTGTGCAATCTAG

AAGTGCTTAGCTTTGCTGATTCTGCCATTGACCTGTTGCCTTCCACAATC
GGAATTTTGAAGAAGCTAAGGCTACTGGATTGACAAATTGTTATGGTCT
TTGTATAGCTAATGGTGTCTTTAAAAAATTGGTCAAACCTGAAGAGCTCT
ATATGACAGTGGTTAATGGAGGAGTTCGAAAGGCGATCAGCCTCACTGAG
5 GATAACTGCAATGAGATGGCAGAACGTTCAAAAGACCTTTCTGCATTAGA
ACTTGAGTTCTTTGAAAACAATGCTCAGCCAAAGAATATGTCATTTGAGA
AGCTACAACGATTCCAGATCTCAGTGGGGTGCTATTTATATGGAGCTTCC
ATAAAGAGCAGGCACTCGTATGAAAACACATTGAAGTTGGTTATTGACAA
AGGTGAATTATTTGAATCTTGAATGAACGGCCTGTTTAAGAAAACAGAGG
10 TGTATGTTTAAGTGTGGGAGATATGAATGATCTTGAAGATRTTGAGGTT
AAGTCATCCTCACAACYTCTTCAATCTTCTTCGTTCAACAATTTAAGAGT
CCTTGTCGTTTCAAAGTGTGCAGAGTTGAAACACTTCTTCACACCTGGTG
TTGCAAACACTTTAAAAAAGCTTGAGCATCTTGAAGTTTACAAATGTGAT
AATATGGAAGAACTCATACGTAGCAGGGGTAGTGAAGAAGAGACGATTAC
15 ATCCCCCAAGCTGAAGTTTTTATCTTTGTGTGGGCTACCAAAGCTATCGG
GTTTGTGCGATAATGTCAAAATAATTGAGCTACCACAACTCATGGAGTTG
GAACTTGACGACATTCCAGGTTTCACAAGCATATATCCCATGAAAAAGTT
TGAAACATTTAGTTTGTTGAAGGAAGAGGTAATATAAATTTTAAATGCT
AATACATTACAAAGGATCTTTTCAGTTAAATCTTTCAAAATATATTGTAA
20 TTTGATTGTATGGGGTATTATTGTTGGATGGGACTATTAATAAATGATTA
TCTTGCAAGTTCTGATTCTTAAGTTAGAGAACTGCATGTTAGTAGTATG
TGGAATCTGAAGGAGATATGGCCTTGCGAATTTAATATGAGTGAGGAAGT
TAAGTTCAGAGAGATTAAAGTGAGTAACTGTGATAAGCTTGTGAATTTGT
TTCCGCACAAGCCCATATCTCTGCTGCGTCATCTTGAAGAGCTTAAAGTC
25 AAGAATTGTGGTTCCATTGAATCGTTATTCAACATCCATTGATTGTGC
TGGTGCAACTGGAGATGAATACAACAACAGTGGTGTAAGAATTATTAAG
TGATCAGTTGTGATAAGCTTGTGAATCTCTTTCCACACAATCCCATGTCT
ATACTGCATCATCTTGAAGAGCTTGAAGTCGAGAATTGTGGTTCCATTGA
ATCGTTATTCAACATTGACTTGGATTGTGCTGGTGCAATTGGGCAAGAAG
30 ACAACAGAAGCAGCTTAAGAAACATCAAAGTGGAGAATTTAGGGAAGCTA
AGAGAGGTGTGGAGGATAAAAGGTGGAGATAACTCTCGTCCCCCTTGTTCA
TGGCTTTCAATCTGTTGAAAGCATAAGGGTTACAAAATGTAAGAGGTTTA
GAAATGTATTACACCTACCACCACAAATTTTAATCTGGGGGCACTTTTG
GAGATTTCAATAGATGACTGCGGAGAAAACAGGGAAAATGACGAATCGGA
35 AGAGAGTAGCCATGAGCAAGAGCAGGTAAGGATTTCAATTTCACTTTCKT
ACTTAATTAATGATTAAGCTCCTGCTTTTTRAATAAAAAAGGGACAAACC
ATTTTCATGACTTAATGTAGCAATACAAGTCATGTATAAGAGTGACCAACT
CTTTTTTATTTATAAAATGACTACAAAATATTTTTTTTTCATTAGAGATCA
TGTATAAATGTGACTAATTTTTTCATCACCTAACTTTAGTTGATAAATCTT
40 TATAAATGTCACTAGTTACTTTTCAGTAAAATAACAAATTTAATAAATTA
TCAACAAAAAGCATCAACTAAAAAAATCCACAAACCGTAATAATTTAAA
ATAAAAGGATTTAACATCTAATACGAACAATTTTTTTTCTAAACATGATT
TGGACCAAATATCACCAGCAACTCAAGTTTGGAATCGATTGAGCTTAAAA

CTTGACCARCATAATTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATAT
AAGTTCGTTTCATCTTTTTCTTGATCTTGATAGCAAGTTGAATSATTTT
CTTCTTCAAAATTGATAAAAATCTACATTATAAAGAGACTAGCTTGAAAA
AAAATGGTCTAGGTGGGTCTGGGTCTGGTAGATGAAGATGGAAGGGAGA
5 GTAGATTTCAAAGACACAAACACATCTTCATTTTATTTATTTATTTATTA
TTATTATTTTTTGATATCTTGCTCATATTTGTTACAGATATGTGAGGTCT
ATTAATCTTTTTAAATATATAAAAAATAAATACATAAATGAGAAAATTAA
ATAAAGAATAAATTAATAAGGGCACAATAGTCTTTTTTGGTAAGACAAGG
ACCAAAAGCGCAACAAAAGTAAACAGTAGGGACCATCCGATTTAAAAAAT
10 TAATTAGGGACCAAAAACATAAATTCACCAAAACCATAGGGACCATTCGT
GTAATTTACTCTTGCTTTTCGTTTTGTTTCATATTTGGGTAACATTTTTT
TTGTACATATCTAGGTAACGAACTTGTTGAAAGTGTTACATCTACGATG
TGACCTACTACAACCGATCATAATGGTCATATATGAACACTTCCAACAAG
TTTGTATCTAGGTGTGTACAAAAAACGATAGTTACCATGATGTGAACA
15 TACCAAAAAATTAATTACCTTAGCAAGTTATTTTCCCATTTAGGTTGTAT
GGAAACAGTTCCTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAA
CTTAACCCTTCAATTAACCTACCTTTTTCTTATTAACCTCAATTTCAAGCT
AAATTCTGATTCTTGTTGAAAGTAAGTTGCATCTTTATGTTTGTATTAT
CTTGTTGCATAGGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAA
20 AGATCCAACCTATTTTTAATCTGTTGGCATTTCATCATTTGCAACTGTT
TCTTGAAAAAAA::TACCTAAAATCAAAATAACCATTTTCATATCCAAAA
TTATAAGAGAGAATTGTTAACGGACATGGAATCATAAATCATTAACACAG
TTCAGTACACAGGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTC
TATCAGAGAAAGAGACATTACAAGAAGCCACTGGCAGTATTTCAAATATT
25 GTATTCCCATCCTGTCTCATGCACTCTTTTCATAACCTCCATAAACTTAA
CTTGAACAGAGTTGAAGGAGTGGAGGTGGTGTGTTGAGATAGAGAGTGAGA
GTCCAACAAGTAGAGAATTGGTAACAACTCACCATAACCAACAACAACCT
ATTATACTTCCCAACCTCCAGGAATTGATTCTATGGAATATGGACAACAT
GAGTCATGTGTGGAAGTGCGGCAACTGGAATAAATTCTTCACTCTTCCAA
30 AAGAACAATCAGAATCCCCATTCCACAACCTCAGTAACATACATATTTAT
GAATGCAAAAGCATTAAAGTACTTGTTTTACCTCTCATGGCAGAACTTCT
TTCCAACCTAAAGCATATCGAGATAAGAGAGTGTGATGGTATTGAAGAAG
TTGTTTCAAAAAGAGATGGTGAGGATGAAGACATGACTACATCTAC:::
:::GCACACAACCACCACTTTTTCCCTCATCTTGATTCTCTCACTCTAAA
35 GCAACTGAAGAATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGG
GGAGCAATGAAATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTT
GATCAATTTGAGGTATGCTTTGTACATATTCAATTATTTATTTAATTTCC
TTGTTAATTTCTTTTTCTTTGCAATATTCTATGAAAAAATCACCAAA
TCACAAATAAGAGATTTAACTTTTATTTACACCCATGCGGACTCAAGA
40 ATGGGATTTGGAGGCATATAAAGTTACATTCATTTGAACAAGTATTACCA
TTT.ATTTGTTATTTATCATTTTCATATCATTTACTGATAACATTTCTTTT
TTACTTTTCTAATTAGAAAAGGTCCACATGTCTAATTAGGTTTTCCATTC
TATGTGAATCCTCTATTCTGTCTGTAATCAAGCATCTTAGATTATTTATC

CATTTTCATAATTGTGTTTATATTGACAGTTTTTTTCTTTTATAGTTGT
AATTGCAACCTGTCATATWTTMWWKKCWWATKYWMWWARTAATACATTT
TATACCCWCTATACTAAGATA

5 **RG2N deduced polypeptide sequence (SEQ ID NO:117)**

LGKTTMMHRLKKVVKEKKMFNFIVEAVVGEKTDPIAIQSAVADYLGIELNEKTKPA
RTEKLRKWFVDNSAGKKILVILDDVWQFVDLNDIGLSPLPNQGVDFKVLTSRDKD
VCTEMGAEVNSTFNVKMLIETEAQSLFHQFVEISDDVDRELHNIGVNIVRKCGGLPI
VIKTMACTLRGKSKDAWKNALLRLVNYNIENIVNGVFKMSYDNLQDEETKSTFLL
10 CGMFPEDFNIPTEELVRYGWGLKLFKKVYTIGEARIRLNTCIERLIHTNLLIEVDDVR
CIKMHDLVRAFLDMYSKVEHASIVNHGNTLEWHVDNMHNSCKRLSLTCKGMSK
FPTDLKFPNLSILKLMHEDISLRFKPNFYEEMEKLEVISYDKMKYPLLSPSPQCSVNL
CVFHLHKCSLVMFDCSCIGNLSNLEVLFSFADSAIDLLPSTIGILKKLRLDLTNCYGL
CIANGVFKKLVKLEELYMTVVNNGGVRKAISL

15

RG2O polynucleotide sequence (SEQ ID NO:118)

TTGTAAAACGACGGCCAGTCGAATCGTAACCGTTCGTACGAGAATCGCTG
TCCTCTCCTTCATTTGAATCATGATATTTGAATATCGATACTTTTGACTG
TAGCTTTTGGGTCGATTTTTTAGCAAGATACATAACTGGCCAAACCCATT
20 GGCTATTTTAGCCCAAATATGAAATGGACTGGATTGTTTTTTCCTTTC
TAACACGCACACATCTGGCGATCAGTATCACTCCATTATGAAGACCTAGT
CAAATTCATTAACGTTTCAGTCGTTCTTCAAAGTTTCAAAGTTCCAACCT
CCAACCTCCCTCTTTTTTTTTTCTTTCCTCGATTCTGATTTGAATCCGAT
TCTGCGACGAAGGAGAGCTTGGTCAGAGGGCTGTGATTCTTGAGTCTTGA
25 CCTCCGAATCTAGCTGGATTATTTTCGACACACCAGACCACGTATCAGGT
TGCTCATCCCGAAATACTGCTTTGCAAACCTGTTGTATCATCGCCTAGGAA
ATTAAAGTTTCTTTTTTGGCTCTGTTACTGAATCAGTAGCTTTGCAACTTG
CTCATTATAAGCTGATCCATATTTTACATATCTTTTGAAGAATAATAGGT
ACTGACTTTACCTTTCTGATGAGAGCGATTAAAGAGATACCTCTGTAAAA
30 TCCATTTTTGTGAAGGGATCTGGGTTAGTTTTTAAAGGATTTGCTACAAC
AGTATCCCAACAAACGATCTATTTCCCATTTNACTCATCCGCTCAAGATCT
ATCCACCTTTATATATGTTAATTGGGAGTCTTCCATGGTGCAATGAATCT
AGGATGCATTTAGAAGCCCAATCCATTACAAGTTTTATCCAATTTTCATG
TGACAAGTTGTTGGTTACTATGTAGGTACTTCCACAATTAAGAATTTCCA
35 GCAATGGATGTTGTTAATGCCATTCTTAAACCAGTTGCCGAGACACTTAT
GGAACCTGTAAAGAAACATCTAGGCTACATCATTTCCAGCACAAAACATG
TGAGGGATATGAGTAACAAAATGAGGGAGTTGAACGCTGCAAGACATGCT
GAAGAAGACCACTTGGACAGGAACATAAGAAGTCTGTTGAGATTTCAAA
TCAAGTTAGGAGTTGGTTAGAAGAAGTAGAAAAGATCGATGCAAAAAGTAA
40 AAGCCCTTCCTAGTGATGTCACCGCTTGTGTCAGTCTCAAGATCAAACAT
GAAGTCGGAAGGGAAGCCTTGAAGCTAATTGTGGAGATTGAAAGTGCCAC
AAGACAACACTCTTTGATCACCTGGACTGATCATCCCATTCCTCTGGGAA

AAGTTGATTCCATGAAGGCATCGATGTCCACAGCATCAACCGATTACAAT
GACTTTCAGTCAAGAGAAAAAACTTTTACTCAAGCATTGAAAGCACTTGA
ACCAAACAACGCTTCCCACATGATAGCGTTATGTGGGATGGGTGGAGTGG
GGAAGACCACAATGATGCAAAGACTAAAAAAGTTGCTAAACAAAATAGA
5 ATGTTCAAGTTATATGGTTGAGGCAGTTATAGGGGAAAAGACGGACCCAAT
TGCTATTCAACAAGCTGTAGCGGATTACCTTCGTATAGAGTTAAAAGAAA
GCACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCCAAC
TCTGGAGAAGGTAAGAATAAATTCCTTGTAATACTTGATGACGTCTGGCA
GTCTGTTGATCTAGAAGATATTGGTTTAAAGTCCTTTTCCAAATCAAGGTG
10 TCGACTTCAAGGTCTTATTGACTTCACGAGACGAACATGTTTGCACAGTA
ATGGGAGTTGGATCTAATTCAATTCTTAATGTGGGACTTCTAATAGAAGC
AGAAGCACAAAGTTTGTTCACAAATTTGTAGAACTTCTGAGCCCGAGC
TCCATAAGATAGGAGAAGATATTGTAAGGAAGTGTGCGGTCTACCTATT
GCC.ATCAAAACCATGGCATGTACTCTTAGAAATAAAAGAAAGGATGCTTG
15 GAAGGATGCACTTTCGCGTATAGAGCACTATGACCTTCGCAATGTTGCGC
CTAAAGTCTTTGAAACGAGCTACCACAATCTCCATGACAAAGAGACTAAA
TCAGTGTTTTGTATGTGTGGTTTGTTCGGAAGACTTCAATATTCCTAC
TGAGGAGTTGATGAGGTATGGATGGGGATTAAAGATATTTGATAGAGTCT
ATACATTTATAGAAGCAAGAAACAGGATCAACACCTGCATTGAGCGACTG
20 GTGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGAT
GCATGATCTGGTCCGTGCTTTTGTTTTAGGTATGTATTCTGAAGTAGAGC
ATGCTTCAGTTGTCAACCATGGTAATATACCTGGATGGACTGAAAATGAT
CCG.ACTGACTCTTGTAAGCAATTTTCAATTAACATGCGAGAGTATGTCTGG
AAACATTCCAGGAGACTTCAAGTTTCCAAACCTAACGATTTTGAAACTTA
25 TGC.ATGGAGATAAGTCGCTAAGATTTCCACAAGACTTTTATGAAGGAATG
GAA. AAGCTCCAGGTTATATCATACGATAAAATGAAGTATCCAATGCTTCC
CTTGTCTCCTCAATGCTCCACCAACCTTCGAGTGCTTCATCTCCATGAAT
GTTCAATTAAGATGTTTGATTGCTCTTGTATTGGAAATATGGCGAATGTG
GAAGTGTTGAGCTTTGCTAATTCTGGCATTGAAATGTTACCTTCCACTAT
30 CGGAAATTTAAAGAAGCTAAGGTTACTTGATTAAACAGATTGTCATGGTC
TTCATATAACACACGGTGTCTTTAACAATTTGGTCAAACCTGAAGAGTTG
TAT.ATGGGATTTTCTGATCGACCTGATCAAACCTCGTGGTAATATTAGCAT
GACAGATGTCAGCTACAATGAATTAGCAGAACGTTCAAAAGGCCTTTCTG
CATTAGAGTTCCAGTTCTTTGAAAACAATGCCCAACCAAATAATATGTCG
35 TTTGGGAAACTTAAACGATTCAAGATCTCAATGGGATGCACTTTATATGG
AGGATCAGATTACTTTAAGAAAACGTATGCTGTCCAAAACACATTGAAGT
TGGTTACTAACAAAGGTGAACCTATTGGACTCTAGAATGAACGAGTTGTTT
GTTGAAACAGAAATGCTTTGTTTAAAGTGTTGATGATATGAATGATCTTGG
TGATGTTTGTGTGAAGTCCTCACGTTCTCCTCAACCTTCTGTGTTCAAAA
40 TTCTAAGAGTCTTTGTCGTTTCCAAGTGTTGAGTTGAGATACCTTTTC
ACAATTGGTGTAGCCAAGGATTTGTCAAATCTTGAGCATCTTGAAGTTGA
TTC.ATGTAATAATATGGAACAACCTCATATGTATTGAGAATGCTGGAAAAG
AGACAATTACATTCCTAAAGCTGAAGATTTTATCTTTGAGTGGGCTACCA

AAGCTTTCGGGTTTGTGCCAAAATGTCAACAACTTGAGCTACCACAACT
CATAGAGTTGAAACTTAAGGGCATTCCAGGGTTCACATGCATTTATCCGC
AAAACAAGTTGGAAACATCTAGTTTGTGTAAGGAAGAGGTAGATATATGT
TTTATGTTAATACAAGTTAAAAAATCTTTTAACTAAAAGTTTCAGTATA
5 TATATCTATATGTCTATAATTTGATTATATGATGTATTAGTGTTTGGATG
TGGCTATTAAGGGATGATTATTTTGCAGGTTGTGATTCCTAAGTTGGAGA
CACTTCAAATTGATGAGATGGAGAATTTAAAGGAAATATGGCATTATAAA
GTTAGTAATGGTGAGAGAGTTAAGTTGAGAAAGATTGAAGTGAGTAACTG
TGATAAGCTTGTGAATCTATTTCCACACAACCCCATGTCTCTGCTGCATC
10 ATCTTGAAGAGCTTGAAGTCAAGAAATGTGGTTCCATTGAATCGTTATTC
AACATCGACTTGGATTGTGTTGATGCCATAGGAGAAGAAGACAACATGAG
GAGCTTAAGAAACATTAAAGTGAAGAATTCATGGAAGTTAAGAGAAGTGT
GGTGTATAAAAGGTGAAAATAACTCTTGCCCCCTTGTTTCTGGCTTTCAA
GCTGTTGAAAGCATAAGCATTGAAAGTTGTAAGAGGTTTAGAAATGTATT
15 CACACCTACCACCACCAATTTTAATATGGGGGCACTTTTGGAGATATCAA
TAGATGACTGTGGAGAATACATGGAAAATGAAAAATCGGAAAAGAGTAGC
CAAGAGCAAGAGCAGGTATGGATTTCAATTTCACTTTCTTACTTACTTAA
GGATTAAGCTTCTGTTTTTTTGAATAAAAAAGGGACATCTTCTAATAATG
CACATCTTAAATTA AAAAGTATTTAATTGTTGCATAGCAGCGTATAACAT
20 CTTCTAATAATTTATCTGAAGGTGAAAGATCCAATACTTCTAATTTGTT
AACAAATTTCAATCATTGCAAAATGTTTCTTAAAAAATTAATTACCTGAAA
TCAAAACAATCTTCTTCAAATCCAAAATTATGAGACAGAATTGAGAAGGG
ATGTGAAATTATAAACCATTAACACAATTCCATGCTCACGTTACTAATTA
CATTTCTTGTTGGGATATATATGTACAGACTGATATTTTGTGAGAGGAAG
25 TGA-AATTACAAGAAGTCACTGATACTATTTCTAATGTTGTATTACATCG
TGTCTCATACACTCTTTTATAACAACCTCCGTAAACTCAACTTGGAGAA
GTATGGAGGAGTTGAGGTTGTGTTTGAGATAGAGAGTTCAACAAGTAGAG
AATTGGTAACAACATAACCATAAACAACAACAACAACAACCTATATTT
CCC.AACCTTGAGGAATTATATCTATATTATATGGACAACATGAGTCATGT
30 ATGGAAGTGCAACAACCTGGAATAAATTTTACAACAATCAGAATCCCCAT
TCCACAACCTCACAACCATAACATGTCCGATTGCAAAAGCATTAAAGTAC
TTGTTTTACCTCTCATGGCAGAACTTCTTTCCAACCTAAAGAGAATCAA
TATTGACGAGTGTGATGGTATTGAAGAAATTGTTTCAAAAAGAGATGATG
TGG.ATGAAGAA

35

RG2O deduced polypeptide sequence (SEQ ID NO:119)

MDV V N A I L K P V A E T L M E P V K K H L G Y I S S T K H V R D M S N K M R E L N A A R H A E E D H L D
R N I R T R L E I S N Q V R S W L E E V E K I D A K V K A L P S D V T A C C S L K I K H E V G R E A L K L I V E I E
S A T R Q H S L I T W T D H P I L G K V D S M K A S M S T A S T D Y N D F Q S R E K T F T Q A L K A L E P N N
40 A S H M I A L C G M G G V G K T T M M Q R L K K V A K Q N R M F S Y M V E A V I G E K T D P I A I Q Q A V A
D Y L R I E L K E S T K P A R A D K L R E W F K A N S G E G K N K F L V I L D D V W Q S V D L E D I G L S P F P
N Q G V D F K V L L T S R D E H V C T V M G V G S N S I L N V G L L I E A E A Q S L F Q Q F V E T S E P E L H K I

GEDIVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRIEHYDLRNVAPKVFETSYHN
LHDKETKS VFLMCGLFPEDFNIPTEELMRYGWGLKIFDRVYTFIEARNRINTCIERL
VQTNLLIESDDVGCVKMHDLVRAFVLGMYSEVEHASVVNHGNIPGWTENDPTDSC
KAISLTCEMSGNIPGDFKFPNLTKLKMHGDKSLRFPQDFYEGMEKLQVISYDKMK
5 YPMLPLSPQCSTNLRVLHLHECSLKMFDCCSIGNMANVEVLSFANSIEMIPLSTIGN
LKKLRLDLTDCHGLHITHGVFNNLVKLEELYMGFSDRPDQTRGNISMTDVSYNE
LAERSKGLSALEFQFFENNAQPNNMSFGKLKRFKISMGCTLYGGSDYFKKTYAVQ
NTLKLVTNKGELLDSRMNELFVETEMLCLSVDDMNDLGDVCVKSSRSPQSPSVFKIL
RVFVVSKCVELRYLFTIGVAKDLSNLEHLEVDSNNMEQLICIENAGKETITFLKIKI
10 LSLSGLPKLSGLCQNVNKLLELPQLIELKLKGIPGFTCIYPQNKLETSSLLKEEVVIPKL
ETLQIDEMENLKEIWHYKVSNGERVKLRKIEVSNC DKLVNLFPHNPMSLLHHLEEL
EVKKCGSIESLFNIDLDCVDAIGEEDNMRS LRNIKVKNSWKLREVWC IKGENN SCPL
VSGFQAVESISIESCKRFRNVFTPTTTNFMGALLEISIDDCGEYMENEKSEKSSQEQ
EQTDILSEEVKLQEVTDITISNVVFTSLIHSFYNNLRKLNLEKYGGVEVVFEIESSTS
15 RELVTYHKQQQQQPIFPNLEELYLYMDNMSHVWKCNNWNKFLQQSESPFHN
LTTIHMSDCKSIKYLFSPLMAELLSNLKRINIDECDGI

RG2P polynucleotide sequence (SEQ ID NO:120)

CCCATTGCTATTTCAGGAAGCAGTAGCAGATTACCTCNGTATAGAGCTCAA
20 AGAAAAAACTAAATCNGCAAGAGCTGATATGCTTCGTAAAATGTTAGTTG
CCAAGTCCGATGGTGGTAAAAATAAGTTCCTAGTAATACTTGACGATGTA
TGGCAGTTTGTGATTTAGAAGATATCGGTTTAAGTCCTTTGCCAAATCA
AGGTGTTAACTTCAAGGTCTTGCTAACATCACGGGATGTAGATGTTTGCA
CTATGATGGGAGTCGAAGCCAATTC AATTCTCAACATGAAAATCTTACTA
25 GATGAAGAAGCACAAAGTTTGTTCATGGAGTTTGTACAAATTCGAGTGA
TGTTGATCCCAAGCTTCATAAGATAGGAGAAGATATTGTAAGAAAGTGTT
GTGGTTTGCCTATTGCCATCAAAACCATGGCCCTTACTCTTAGAAATAAA
AGCAAGGATGCATGGAGTGATGC ACTTTCTCGTTTAGAGCATCATGACCT
TCACAATTTTGTGAATGAAGTTTTTGG AATTAGCTACGACTATCTTCAAG
30 ACCAGGAGACTAAATATATCTTTTGTCTTTGTGGATTGTTTCCCGAAGAC
TACAATATTCCTCCTGAGGAGTTAATGAGGTATGGATGGGGCTTAAATTT
ATTTAAAAAAGTGTATACTATAAGAGAAGCAAGAGCCAGACTCAACACCT
GCATTGAGCGGCTTATCCATACCAATTTGTTGATGGAAGGAGATGTTGTT
GGGTGTGTAAAGATGCATGATCTAGCACTTGCTTTTGTATGGATATGTT
35 TTCTAAAGTGCAGGATGCTTCAATTGTCAACCATGGTAGCATGTCAGGGT
GGCCTGAAAATGATGTGAGTGGCTCTTGCCAAAGAATTCATTAAACATGC
AAGGGTATGTCTGGGTTTCCTATAGACCTCAACTTTCCAAACCTCACAAT
TTTAAACTTATGCATGGAGATAAGTTTCTCAAGTTTCCTCCAGACTTTT
ATGAACAAATGGAAGCTTCAAGTTGTATCGTTTCATGAAATGAAATAT
40 CCGTTTCTTCCCTCGTCTCCTCAATATTGCTCCACCAACCTTCGAGTTCT
TCATCTCCATCAATGCTCATTGATGTTTGATTGCTCTTGTATTGGAAATC
TGTTTAATCTGGAAGTGTTGAGCTTTGCTAATTCTGGCATTGAATGGTTA

CCTTCCAGAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACAGA
TTGTTTTTGGTCTTCGTATAGATAAGGGTGTCTTAAAAAATTTGGTCAAAC
TTGAAGAGGTTTATATGAGAGTTGCTGTTTGAAGCAAAAAAGCCGAAAT
AGAAAAGCCATTAGCTTCACAGATGATAACTGCAATGAGATGGCAGAGCG
5 TTC

RG2P deduced polypeptide sequence (SEQ ID NO:121)

PIAIQEAVADYL?IELKEKTKSARADMLRKMLVAKSDGGKNKFLVILDDVWQFVDL
EDIGLSPLPNQGVNFKVLLTSRDVDVCTMMGVEANSILNMKILLDEEAQSLFMEFV
10 QISSDVDPKLHKIGEDIVRKCCGLPIAKTMALTLRNKSKDAWSDALSRLHHDLHN
FVNEVFGISYDYLQDQETKYIFLLCGLFPEDYNIPPEELMRYGWGLNLFKKVYTIRE
ARARLNTCIERLIHTNLLMEGDVVGCVKMHDALALAFVMDMFSKVQDASIVNHGS
MSGWPENDVSGSCQRISLTCKGMSGFPIDLNFPNLTILKLMHGDKFLKFPDFYEQ
MEKLQVVSFHEMKYPFLPSSPQYCYSTNLRVLHLHQCSLMFDCSCIGNLNFNLEVLSE
15 ANSGIEWLPSRIGNLKKLRLLDLTDCFGLRIDKGVKLNLVKLEEVYMRVAVRSKKA
GNRKAISFTDDNCNEMAERS

RG2Q polynucleotide sequence (SEQ ID NO:122)

TGGGGAAGACACAGTGATAGAAAAAAGAAATGTTGTGGAAAAGAGGA
20 AAATGTTTGATTATGCTGTTGTGGCGGTTATAGGGGAAAAGACGGACCCT
ATTGCTCTTCAGAAAAGTGTGCGGATTACTTGCATATTGAGCTAAATGA
AAGCACTAACTAGCAAGAGCAGATAAACTTTGCAAATGGTTCAAGGACA
ACTCGGATGGAGGTAAGAAAAAGTTCCTCGTAATACTCGACGATGTTTGG
CAATCTGTTGATTTGGAAGATATTGGTTTAAGTACTCCTTTTCCAAATCA
25 AGGTGTCAACTTCAAGGTTTTGTTGACATCACGAAAGAGAGAAATTTGCA
CAATGATGGGAGTTGAAGCTGATTTAATTCTCAATGTCAAAGTCTTAGAA
GAAGAAGAAGCACAAAAGTTGTTCCCTCCAGTTTGTAGAAATTGGTGACCA
ATACCACGAGCTTCATCAGATAGGGGTACATATAGTAAAGAAGTGTTATG
GTTTACCCATTGCCATTAAACCATTGGCTCTTACTTTAAGAAATAAAAGA
30 AAGGATTCATGGAAGGACGCACTCTCTCGTTTAGAGGACCATGACACTGA
AAATGTTGCAAATGCAGTTTTTCGAGATGAACTACCGCAATCTACAAGATG
AGGAGACCAAAGCCATTTTTTTGCTTTGCGGTTTGTTCCTCCGAAGACTTT
GAT.ATTCCTACTGAGGAGTTGGTGAGGTATGGATGGGGCTTAAATCTATT
TAAAAAAGTGTATACCATAAGAAAGGCAAGAACGAGATCGCATACATGTA
35 TTGAGCGACTCTTGATTCAAATTTGTTGATTGAAAGTAACGATATTCGG
TGCGTCAAGATACACGATCTGGTGCGCGCTTTTGTGTTTGGATATGTATTG
TAAAGTTGAGCATGCTTCAATTGTCAACCATGGTAATATGCGGACCGAAT
ATAATATGGCTGACTCTTGCAAAACAATTCATTAAACATACAAGAGTATG
TCTGGGTTTGAAGTTTCCAGGAGACCTCAAGTTTCCAAACCTAACAGTTTT
40 GAACTTATGCANGGAGATAAGTCTCTAAGGTTTCCTCAAGACTTTTATC
AATCAATGGAAAAAAGTTCGGGTTATATCATATGATAAAATGAAGTATCCA
TTGCTTCCCTCATCACCTCAATGCTCCACTAACATCCGAGTGCTTCGTCT

CCATGAATGTTTCATTAAGGATGTTTGATTGCTCTTGTATTGGAAAGCTAT
TGAATTTGGGAAGTCCTCAGCTTTTTTAATTCTAACATTGAATGGTTACCT
TCCACAATCAGAAATTTAAAAAAGCTAAGGCTACTAGATTTGAGATATTG
TGATCGTCTTCGTATAGAACAAGGTGTCTTGAAAAATTTGGTCAAACCTTG
5 AAGAACTTTATACTGGATATACATCAGCGTTTACAGA

RG2Q deduced polypeptide sequence (SEQ ID NO:123)

GEDTVIEKKKNVVEKRKMFDYAVVAVIGEKTDPIALQKTVADYLHIELNESTKLAR
ADKLCKWFKDNSDGGKKKFLVILDDVWQSVLEDIGLSTPFPNQGVNFKVLLTSR
10 KREICTMMGV EADLILNVKVL EEEEEAQKLF LQFVEIGDQYHELHQIGVHIVKKCYG
LPIAIKTMALTLRNKRKDSWKDALSRLEDHDTENVANAVFEMNYRNLD EETKAI
FLLCGLFPEDFDIPTEELVRYGWGLNLFKKVYTIRKARTRSHTCIERLLDSNLLIESN
DIRCVKIHDLVRAFVLDMYCKVEHASIVNHGNMRTEYNMADSKTISLTYKSMMSG
FEFPGDLKFPNLTVLKLM?GDKSLRFPQDFYQSMEKLRVISYDKMKYPLLPSSPQCS
15 TNIRVLRRLHECSLRMFDCSCIGKLLNLEVL SFFNSNIEWLPSTIRNLKKLRLDLLRYC
DRLRIEQGV LKNLVKLEELYTGYTSAFTE

RG2S polynucleotide sequence (SEQ ID NO:124)

ATTTGGGGTTTTACATTTAATTTTTTGTGCATGAATGTGAAAATAGACTG
20 CTTATTGATTCTTTGTGTTTCATTGAGTTGATTTTCATTATTACTACCTT
ACAAATTGCTCAGTGATAGATTTCCATTAATTTGCTAATTCGGTTGCTTC
TAAATATGTAGGAGCTACTAAAAGCAAAAATATCGAGCAATGTCGGACCC
AACGGGGATTGCTGGTGCCATTATTAACCCAATTGCTCAGAGGGCCTTGG
TTCCCGTTACAGACCATGTAGGCTACATGATTTCTGCAGAAAATATGTG
25 AGGGTCATGCAGACGAAAATGACAGAGTTGAATACCTCAAGAATCAGTGT
AGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCAGATTCCATCTC
AAATTAAGGATTGGTTGGACCAAGTAGAAGGGATCAGAGCAAATGTGGAA
AACTTTCCGATTGATGTCATCACTTGTTGTAGTCTCAGGATCAGGCACAA
GCTTGGACAGAAAGCCTTCAAGATAACTGAGCAGATTGAAAGTCTAACAA
30 GACAGCTCTCCCTGATCAGTTGGACTGATGATCCAGTTCCTCTAGGAAGA
GTTGGTTCCATGAATGCATCCACCTCTGCATCATCAAGTGATGATTTCCC
ATCAAGAGAGAAAACCTTTACACAAGCACTAAAAGCACTCGAACCCAACC
AACAAATCCACATGGTAGCCTTGTGTGGGATGGGTGGAGTAGGGAAGACT
AGAATGATGCAAAGGCTGAAGAAGGCCGCTGAAGAAAAGAAATTGTTTAA
35 TTATATTGTTAGGGCAGTTATAGGGGAAAAGACGGACCCCTTTGCCATT
AAGAAGCTATAGCAGATTACCTCGGTATACAACCTCAATGAAAAAACTAAG
CCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAAAAGAATTCAGATGG
AGGTAAGACTAAGTTCCTCATAGTACTTGACGATGTTTGGCAATTAGTTG
ATCTTGAAGATATTGGGTAAAGTCCTTTTCCAAATCAAGGTGTCGACTTC
40 AAGGTCTTGTGACATCACGAGACTCACAAGTTTGCACTATGATGGGGGT
TGAAGCTAATTCAATTATTAACGTGGGCCTTCTAACTGAAGCAGAAGCTC
AAAGTCTGTTCCAGCAATTTGTAGAACTTCTGAGCCCGAGCTCCAGAAG

ATAGGAGAGGATATCGTAAGGAAGTGTTGCGGTCTACCTATTGCCATAAA
AACCATGGCATGTACTCTTAGAAATAAAAGAAAGGATGCATGGAAGGATG
CACTTTCGCGCATAGAGCACTATGACATTCACAATGTTGCGCCCAAAGTC
TTTGAACGAGCTACCACAATCTCCAAGAAGAGGAGACTAAATCCACTTT
5 TTTAATGTGTGGTTTGTTCCTCGAAGACTTCGATATTCCTACTGAGGAGT
TGATGAGGTATGGATGGGGCTTGAAGCTATTTGATAGAGTTTATACGATT
AGAGAAGCAAGAACCAGGCTCAACACCTGCATTGAGCGACTGGTGCAGAC
AAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATGCATGATC
TGGTCCGTGCTTTTGTTCCTGGGTATGTTTTCTGAAGTCGAGCATGCTTCT
10 ATTGTCAACCATGGTAATATGCCCCGAGTGGACTGAAAATGATATAACTGA
CTCTTGCAAAAGAATTCATTAACATGCAAGAGTATGTCTAAGTTTCCAG
GAGATTTCAAGTTTCCAAACCTAATGATTTTGAAACTTATGCATGGAGAT
AAGTCGCTAAGGTTTCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCA
TGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTC
15 GATGCTCCACCAACATTCGGGTGCTTCATCTCACTAAATGTTCAATAAG
ATGTTTGATTGCTCTTGTATTGGAAATCTATCGAATCTGGAAGTGCTGAG
CTTTGCTAATTCTCGCATTGAATGGTTACCTTCCACAGTCAGAAATTTAA
AGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAA
CAGGGTGTCTTGAAAAGTTTAGTCAAACCTTGAAGAATTTTATATTGGAAA
20 TGCATCTGGGTTTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTG
ACAACCTTTCTGCATTAGAATTCGCGTTCCTTAATAACAAGGCTGAAGTG
AAAAATATGTCATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGACG
CTCTTTTGATGGAAATATCAATATGAGTAGCCACTCATACGAAAACATGT
TGCAATTGGTGACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGG
25 TTATTTTGAACAAAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGA
TCTTGAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCAT
TCTGCAATTTAAAAGTTCTTATTATTTCAAAGTGTGTAGAGTTGAGATAC
CTTTTCAAACCTCAATCTTGCAAACACTTTGTCAAGACTTGAGCATCTAGA
AGTTTGTGAATGCGAGAATATGGAAGAACTCATACATACTGGAATTTGTG
30 GAGAAGAGACAATTACTTTCCCTAAGCTGAAGTTTTATCTTTGAGTCAA
CTACCGAAGTTATCAAGTTTGTGCCATAATGTCAACATAATTGGGCTACC
ACATCTCGTAGACTTGATACTTAAGGGCATTCCAGGTTTCACAGTCATTT
ATCCGCAGAACAAAGTTGCGAACATCTAGTTTGTGTAAGGAAGAGGTAGAT
ATATGTTCTTTATGTTAATAACAATTTAAATAATATTTTCAACCAAATTT
35 CATAATATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGG
CTATTAAGGGATGATTATTTTGCAGGTTGTGATTCCTAAGTTGGAGACAC
TTC.AAATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTT
AGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGA
TAAGCTTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATC
40 TTGAAGAGCTTAAAGTCAAGAATTGCGGTTCCATTGAATCGTTATTCAAC
ATTGACTTGGATTGTGTGCGGTGCAATTGGAGAAGAAGACAACAAGAGCCT
CTT.AAGAAGCATCAACATGGAGAATTTAGGGGAAGCTAAGAGAGGTGTGGA
GGATAAAAGGTGCAGATAACTCTCATCTCATCAACGGTTTTCAAGCTGTT

GAAAGCATAAAGATTGAAAAATGTAAGAGGTTTAGCAATATATTCACACC
TATCACCGCCAATTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAG
GTTGCGGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTTC AATTTAA
CTTTCTTAAGTAATTAAGGACTAACCTCCTGTTTTTTGAATAATAAAGAG
5 GTGGGATGACTAAACTTGGGCATCACAATTGCAACAAAATGTTACAAACC
ATGAAACGTTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTT
AAAAATATGGCTTAAATTTTTTTATATTTATGTATCAACATGATTTTTCA
TTAGAGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATTAGAACT
AATTCTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTT
10 AGTGAAATAAGCAACGGATTTAATAAGTTAACAACCTAAATGTCATTTCC
TAACAAAAAAACTATTTGGTTCAGAAGAACCGTAATTC AAGATAACTAA
AATAAAAAATATTTGACATTCACTAAGAGCATTTTTTTTTCTAAATATGAT
TGCAAATGAATAAAACTTAAATTTATACAGAAAAGATTTTTATATATGTT
ATACAAAATTTACAAATTGAACTGGATATGTTAATTAACGGTTTATAAT
15 TCTGGTATCACAAAGGGATATATAAAAAATATTATTTCTGTAGTCATT
TATAATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAG
TTAACTTTTCATAATAAAAAATTATAATTATTATTTATTTTAAATAAATT
ATT.AATTAAGAGATGTATCAAAAATTTAAAGTTATTATAACTTCAAATTT
AACATATAATTAGAAAATATATGATCATAACTTTCCGCAACTCTTCTTTT
20 GTATTAAAATGCCCAGAGAAGCTCTTAGTAYATTTTCTAAATCAAAGTCA
CAAACTAATGAAGCATATAATTTTGTGAAAATCAATTAGCATTAGGTTT
TAAGAGTCAACCAAATTC AAAGAGTAATCCAATGCTTTCATTACCACTATG
GAGAAAATATTTTCTTAGTTTAAATGAAATGAAAACAAACATTC AAACTA
ATTGTTGCTTACTAAACCAAAGACCCATTACTTAGCCAAGAGTTTAACCA
25 AAAAAAATTACATTCATGTATCATTATTCATGACTAGATATATATGAACA
TGAAGGGAGTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTC
ATGGAATTCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCA
ACC.AAAGAGAAGTTTAGCCTAGCATGGCTAAACTCAAGAAAATAAAAAATA
GGATTAGAAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTT
30 GTTCTTGATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAAT
GCAAAGGAGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAGCTGCA
CCAACCTCCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCTG
AAGGACCCTATGCGGGTGCCTTGCGCGGGTGGAGCTGAATACGAAAGGTC
TTTGGTCTTTGTGAGGGTGATGCTGTGCGGGTTAGCTTGTCGCATGCTTC
35 CGCGCGGTTGCGGCACATGTGCACAAGTGATGCATGGTGTGTACGTTCTT
GAGTTTTGAGCCTCCGATGCTTAGTCCATTTGGCCCAATTCGAGTCCAAT
CAGCTTATGACCCATTTTTCTTCAAGTTATCTTCAAGTTATCTTCAAGTT
AAGCCCAAATTGCCTTCTCCAAATCATCCATAACTTCACAAAATCGCCCG
TTCATCTTAATCCCGAATGCACAATTATTCTCCTGTCTTCCTTTTAAGCA
40 AGATACCACCTTCTTCATGCTTCATCCATCAATAGTACACTTCATGTATC
ATCTCTACTAGTTATTTAGTCCACAATCCTTATTGTCCTCCAAATTTAAT
TATCTCATTTAGTTCCCGTTCCACTAGTTTCCTTAAATTTGCAATTAAG
CTC.ACACAAATATTAAGTACCTGAAATGGTCATAAAATAACAAAAGGAA

AATATGCATGAAGATTAACATAAATGATGAACGAAATATGCTAAAAATAGAC
TATAAAATGAAGTAAATAAAATGAAATTATCGCACTCCGACCACCCTTAT
AGGCTTGTAGTCCATCCACCCTTCATTCTTGTACCAATATGGGATGGAA
ACATCATTAAATTAAGCCAAAAAACTAACATATAAGGGGTGAGTGACAAAG
5 GTAAGTACTAAAGATGAAAATAATCCATTTTTYTTGTATATACACAACAC
ACACATAGGGGCAGACGTAGGATTTTCATAGTACAGATTGTTGGTGGCACA
TAAGTGTTGCTGGTGACACTTTTTTTTTCTTTTACGTAGTGGCACAACAG
TAGAAAAAACGARAAATTCGAAATTTTTTACAATGTGTSTAAAAAAAAYA
GTGGTTGTTGGTGCCACTATGGACACCAAAGTTGAACTGCCCCTGCGCGC
10 RCACACACACACACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
ARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTTAGGAA
AATGGAGTTATATCTTTGATATTGTATTTTTTTAATGTAATTTATATATT
TAATCATTTTTAGTTTATAAGTTTTATTTATTTTGATATGAAAAAAAAGT
CTTTTATACATTGGATTTAACATAAAAATCCAACAATATTAATCAAAAAG
15 ACCAMACATGTGGACAMWTATGTATATAAWTAATTCACAATAGTCTTTAG
GAATAGNATTATATATATAATTAATTCTCAATGGTCTTAGGAATAGTAAG
TTCTTATATTTCAAACCTTNGCCACAATTCTTTGKTTACTTWGACACTTY
CCTCTCTCTAATTATATATATATATATATATATATATATATATATACACA
CACACACACACACACTAGATGTGTGCCCGCGCAAAGCAGTGACGTNNNGG
20 AGAANACTTTCTTAAGCATAAATAATTATTATATTTTTTTATTGGGTATTA
TATAATAAAAAATTACAACCTTTTAAATAAAAATTTTATGTTTATACTTTA
TATTTATATTGCTTGTATACTATTAATATAATAAATTAATATTTATGTCT
AATTTATGAAATGTAAATTAATTTAAATACATGAATTTAATATTTTTAAA
ATTTTCAGTTTGCTTCAAATTGAGTTTCTTAATTATTTTTTTTAATTCAN
25 GTATTCAAACCTTTGGTAAGTATTAAGAATTATTTATGCACAATTGATT
TATACAAAAAATTTGTAACCTTATACATCTTAAAATTCAAGATATAACTA
ACATGTTTTACAATATATATATATATATATATATATATATATATATATAT
ATATATATATATATATAGTAAAGCGCANAGGTCATAGGNANAGANTATTT
TCTATTATTCTACGTTTTGCCACAAAAGTTTGAACACTTTGCCACTTTTT
30 GTCCCTCCTTAACCTTTTCAATGTTTTGCGACAAAAGTTCCAAAACCTTG
CCACTTTGATCATTCTCAACTTTTCACCGCATTAGTTTGTGGAGTTGGC
AGTTTTGGTCCCCCTAACTTCGATATTTTCTCCTGCTAGCCAAAAAGGGT
TCCAGAGTTTCACANTTTTGGTCCCTGACAATAACCAAATGTGAGATGTC
AAATTTTTGCCACATTAGTTTGTGGAGTTGTCCCTTTTGGTCCCCCACA
35 TTCGATATTCTACTATACGACCTTATTTTTCTCAAATAACAACACGTATA
TTTAATTACCAATGATAGAAATAGATATCAAATAAAGTATTTGTAACACC
GTGTAAGAACGGTGCTACTATAGGTAAAAATAAACATTTCAAAGTACGAT
GTCCTAATTGGAAAAAGAGTTTTAAAAAAAATAACAACCTAGGGGCGAGTTT
TTTTTACAAGTTTGTATCAAATCATATCAAATTTAAGGTGGAACGGTGA
40 CCACATTAACCAGAAATGTAATTTATCTTTGATTTTGATAATTTTTAAT
ATTTTGTGTGATCTATGTATTTAAAAGTAAACAACAAAGAACATAATCC
AAAACCCTAAATTGCAAGTCTCGCCCAATTTCTCTATCACTAGTCGTCAC
TTACGATGGCGTTACGTCGCTCTCTCACTTCTTACAACCCTTTGTGCTA

CTCATTACAATAACGAAAAGTTGAATATCCATATATTTATTTGGATGTGG
AATTGAACAAATCTCGTCAAATTTTTGATTTTGTGATGGATTGAGTAG
AAGTTTGGGCAGAACGGGAATGATGGTCTGCAAGTGGTTATAAACTTGAT
TCTGAGTTATTACTATATATGTAGCCTCTTTACAACGACCAAGGTTTCTT
5 CCAGGTACCATTTGATCTTTTTAGAACCCAGTTGTCTGAAACACCCTGAT
TTGGATCAAATATCACCAACAACCTCTTAAAACTTGATTAATCAATTGTT
TTCTTCATCTTGATAACAAGTGAATGATTTTCTACTTAGATTAACCTGA
AAAAAAAGGTCCATGTGCGTCTGGTGGATCTGGTAAATGAAGATGGAAGG
GAGAGCTGACTTTAAAGACACAAACACGTCACCATATCTTTTATTTTATT
10 TTAAATTTGCTTTTTTCCTATTTCTTTCTTTCTTGATCTCCAGATGGTAT
GTGGTGTGGATAATTTACACATAGAGATTGGGAACGACTGTGTTTTAGAG
AGGACGTGGCTTGGGGTTGAGGATGGTTTATGGCTGGCCGAGTTTCATTT
ATATAAACAAACAAATATATAAAACAAGGGGTAAAATGGCCATCTTATAT
GTATTTAACCGTCCTTTTTTATTTTTTTTTTATTTTTAAATTTAAGAAGG
15 GGTATACCAGTGTGAGCCTCTTATCCCAACCAGGCAACCAGTCAAATAG
GGACTTAGGTTGTTTGGAAACAGTTCCTGTGAGACCGTGACTTGGATGGTA
GATAAATTTAGTAACTTAACCCTTCAATTAACCTACCTTTTTCTTATTA
ACTCAATTTCAACCTAAATTCTGATTCTTGTTTGAAAATAAGTTGCATCT
TTATGTTTGTATTATCCTGTTGCATAGGATCCTTAGCATCTTTTAATAAT
20 TTATTTGAAGGTGAAAGATCCAACCTATTTTTTAGCTGTTGGCATTTTCCA
TCATTTGCAACTGTTTCTTGAAAAAAAATACCTAAAATCAAAATAACCA
TTTTCAAATCCAAAATTATAAGAGAGAATTGTTAATGGACGTGGAATCGT
AAATCATTAACACAGTTCAGTACACAAGTTGCTAATTACATTTCTTGCTG
TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGTCACTGAT
25 ACTAATATTTCTAATGATGTTGTATTATTCCTATCCTGTCTCATGCACTC
TTTTCATAACCTCCATAAACTTAAATTGGAGAGAGTTAAAGGAGTGGAGG
TGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTGGTAACA
ACTCACCATAACCAACAACATCCTATTATACTTCCCAACCTCCAGGAATT
GGATCTAAGTTTTATGGACAACATGAGTCATGTGTGGAAGTGCAGCAACT
30 GGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCAC
AACCTCACAACCATAACATGTTTCACTGTCAGAAGCATTAAGTACTTGTT
TTCGCCTCTCATGGCAGAACTTCTTTCCAACCTAAAGGATATCTGGATAA
GTGGGTGTAATGGTATTAAAGAAGTTGTTTCAAAGAGAGATGATGAGGAT
GAAGAAATGACTACATTTACATCTACCCACACAACCACCATCTTGTTCCC
35 TCATCTTGATTCTCTCACTCTAAGACTACTGGAGAATCTGAAGTGTATTG
GTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAAATATCTTTCAATAAT
ACCCTGCAACTACTGCTGTTCTTGATCAATTTGAGGTATGCTTTGTACA
TATTCAATTATTTATTTAATTTCTTTTTCTTTGCAATATTCTATAAAT
AATACATTTTATACCCACTATACTAAGATAATAATTACCTAGAGGGATGG
40 ATGCTATGACACAGCTGCTACACTTCAGAACTCTAGTAAGGGCAGTTAT
GGAAGTTCAATAAAATGATAATGGCATCTTTTGATGGGTAATATAGGCAA
TTTAAAGTTTTATTTCTGTTAAAGCAGTATTTAGCAAGTACTGGCCAGTAG
GAGAGGAGAATATCACCTTTTGTGAAAATCTGGTCATTGTACCCAAGAAT

TTAGTTAAATGTAACATTTTAGATATCAGGGGACATCAGGTGACAGATAT
TGTAGAATAGAACAAATATATAATATTACCCAAAACATTTTTTCTAAGGT
TATTCTGTAAATATGTGCTTTCTTGATTTCATTGAATTTGCATTCCCTAT
ATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAATCCCGAAATTAATTA
5 AAAAAAAAAAAACAAAAGTAAATTTTTGATATGGAGAGCACTGGTATCA
TTTAGTATATAAAAAAACTAGATTTTGAATTAAGTTTCTTATATAAAAGC
TGTGTATATAGTTTAATTAGTTTTACATCATTTTTTCCATGTGGTGTGCA
GTTGTCTGAAGCAGGTGGTGTCTTGAGTTTATGCCAATACGCTAGAG
AGATAGAGATATCTAAGTGTAATGTATTGTCAAGTGTGATTCCATGTTAT
10 GCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGAGAGTAACGGGTTGTGA
TGGCATGAAGGAGGTATTTGAACTCAATTAGGGACGAGCAGCAACAAAA
ACAGAAAGGGTGGTGGTGAAGGAAATGGTGGAAATCCAAGAGTAAAT
AACAAATGTTATTATGCTTCCCAATCTAAAGACATTGAAAATCTACATGTG
CGGGGGTTTGGAACATATATTACATTCTCTGCACTTGAAAGCCTGACAC
15 AGCTCCAAGAGTTAAAGATAGTGGGTTGCTACGGAATGAAAGTGATTGTG
AAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACAACAACAAC
AACGAAGGGGGCATCTTCTTCTTCTTCTTCTTCTTCTAAGAAGGTTG
TGGTCTTTCCCGTCTAAAGTCCATTGAACTATTCAATCTACCAGAGCTG
GTAGGATTCTTCTTGGGGATGAATGAGTCCGGTTGCCTTCATTGGAAGA
20 AGTTACCATCAAGTATTGCTCAAAAATGATGGTGTGTTGCAGCTGGTGGGT
CCACAGCTCCCCAACTCAAGTATATACACACAAGATTAGGCAACATACT
CTTGATCAAGAATCTGGCCTTAACCTTCATCAGGTATATATATATTCCTT
TAATTGGCATGATCTAATTAAGAAAGATATCATTCCTGCCAAGTAAATTT
ACTTCAAACACATTACACTGGTTTCAGTCTAAGTTTATGTTGTTCTAGG
25 AAGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTTCAAGTGGA
AAGGGTATTTTAGGTATTTCTGTCAAAAGTTGTTATTGCAGGCTTTTGA
GTACCTGGAATCGTGTGTGGGAGGAGCGTTATTATTCTGATTTGCTTGT
TCTTTATCATTTTTTCTTAGCCTCTCGAACAGCTAGAAACCCTTTTAATC
TTTTGATTTTAAATGACAAAATTTTCCCTGTACTCTATTTGATTGTTG
30 TTCTTCATGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCTTTTGAT
TGTTATTTTCATATCATGTTGTCTTTGAATCAAGCTTTTCCATTTTCAA
CCAGGGCAAAGGTCAAAGTAACCTACTTTATGAGATCAAAAACAGCAA
CCCATCGGATAACTTTTAGTTGGAGTTAATAGTTACAATTACCATTGTGA
TTAATAATTATAATATCTTGTATTAATTCATTAAAAATTGGTACAGCACAT
35 ATATGACATTTTAAAGGTTTGTGTTTGTGTTWGACATATATATGCCTCTGGC
GTTTTCTTTATTGGACATGCAGACCTCATTCCAAAGTTTATACGGTGACA
CCTCGGGCCCTGCTACTTCAGAAGGGACAACCTGGTCTTTTCATAACTTG
ATCGAATTAGATATGGAATTAATTATGATGTTAAAAAGATTATTCCATC
CAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAGATTATGTGAGTAGTT
40 GTTATTGGGTAGAGGAGGTATTTGAACTGCATTGGAAGCAGCAGGGAGA
AATGGAAATAGTGGAATTGGTTTTGATGAATCGTCACAACTACTACTAC
TACTACTCTTTTCAATCTTCGAAACCTCAGAGAAAATGAAGTTGCATTTTC
TACGTGGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTTGAG

TTTCCAAACCTAACAAGAGTTCATATAAGTAGGTGTAGAAGGTTAGAACA
TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAG
ATATTAGTTGGTGCAACCATATGGAGGAGGTGATTGTTAAGGATGCAGAT
GTTTCTGTTGAAGAAGACAAAGAGAGAGAATCTGATGGCAAGACGAATAA
5 GGAGATACTTGTGTTACCTCGTCTAAAATCCTTGAAATTAAAATGCCTTC
CATGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATT
TTGGATACTTTAGAAATCTACAAATGCCAGCAATAACGACCTTCACCAA
GGGAAATTCTGCTACTCCACAGCTAAAAGAAATAGAAACAAGATTTGGCT
CGTTTTATGCAGGGGAAGACATCAACTCCTCTATTATAAAAAGATCAAAC
10 AACAGGTAAATCAGATCTTTGTTGCTTTAATAATTCTTAAACTACATTG
AAAAGCTTCATGCAAGTTTTTTTTGTTATATTGTCAAAAACCGCAACCTA
CATTTTCAGCTTTATATTTATGTACTTTATGCAGGAGTTCAAACAAAAC
CTGATTAATGTGAAGTGAATATTAAGGTAAATTATATTTTCATGTTCT
AGTTGCCTATTAATTAATGGCCTTTTAGTTCRTGATTTTGGATGTAGTY
15 WTCATGATGATGTGAATCTTCTAATACCCCATTCATTGTTTGGTTGAATG
TTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTTCATCATATG
AAGGACATTAAAGAACATGGATGCTATGAAGATGTTGGAARAC

RG2S deduced polypeptide sequence (SEQ ID NO:125)

20 MSDPTGIAGAIINPIAQRALVPVTDHVGYMISCRKYVRVMQTKMTELNTSRISVEEH
ISRNTRNHLQIPSQIKDWLDQVEGIRANVENFPIDVITCCSLRIRHKLQKAFKITEQI
ESLTRQLSLISWTDDPVPLGRVGSMASTASSSDDFPSREKTFTQALKALEPNQQF
HMYALCGMGGVGKTRMMQRLKKAEEKLKNFYIVRAVIGEKTDPFAIQEAIADYL
GIQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGLSPFPNQG
25 VDFKVLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETSEPELQKIGED
IVRKCCGLPIAKTMACTLRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQE
EETKSTFLMCGLFPEDFDIPEELMRYGWGLKLFDRVYTIREARLNTCIERLVQT
NLLIESDDVGCVKMHDLVRAFLGMFSEVEHASIVNHGNMPEWTENDITDSCKRIS
LTCKSMSKFPDGFKFPNLMILKLMHGDKSLRFPQDFYEGMEKLHVISYDKMKYPLL
30 PLAPRCSTNIRVLHLTKCSLKMFDSCIGNLSNLEVLSFANSRIEWLPSTVRNLKKLR
LLDLRFCDGLRIEQGVLSLVKLEEFYIGNASGFIDDNCNEMAERSDNLSALEFAFF
NNKAEVKNMSFENLERFKISVGRSFDGNINMSSHSYENMLQLVTNKGVDLDSKLN
GLFLKTKVLFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLISKVELRYLFKLN
ANTLSRLEHLEVCECENMEELIHTGICGEETITFPKLKFLSLSQLPKLSSLCHNVNIIG
35 LPHLVDLILKGIPGFTVIYPQNKLRSSLLKEEVVIPKLETQIDDMENLEEIWPCELS
GGEKVKLREIKVSSCDKLVNLFPRNPMSLLHHLEELKVKNCGSIESLFNIDLDCVGA
IGEEDNKSLLRSINMENLGKLREVWRIKGADNSHLINGFQAVESIKIECKRFSNIFT
PITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEVTDNINSDVFLFPSCLMH
SFHNLHKLKLERVKGEVVFEIESESPTSRELVTTHHNQQHPIILPNLQELDLSFMD
40 NMSHVWKCSNWNKFFTLPKQQSESPFHNLTTHMFSCRSIKYLFSPMAELLSNLK
DIWISGCNGIKEVSKRDEDEEMTTFTSTHTTTILFPHLDSLTLRLLENLKCIGGGG
AKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIEISKCNVLSSVIPCY

AAGQMQLQVLRVTGCDGMKEVFETQLGTSSNKNRKGGGDEGNNGGIPRVNNNVI
MLPNLTKLKIYMC GGLEHIFTSALES LTQLQELKIVGCYGMKVIVKKEE DEYGEQ
QTTTTTTTKGASSSSSSSSKKVVVFPRLKSIELFNLPELVGFFLGMNEFRLPSLEEV
IKYCSKMMVFAAGGSTAPQLKYIHTRLGKHTLDQESGLNFHQTSFQSLYGDTSGPA
5 TSEGTTWSFHNLIELDMELNYDVKKIIPSELLQLQKLEKIHVSSCYWVEEVFETAL
EAAGRNGNSGIGFDESSQTTTTTTLFNLRLNREMKLHFLRGLRYWKS NQWTA FEF
PNLTRVHISRCRRLEHVFTSSMVGSLLQLQELDISWCNHMEEVIVKDADVSVEEDK
ERESDGKTNKEILVLPRLKSLKLKCLPCLKGFSLGKEDFSFPLD TLEIYKCPA ITTFT
KGN SATPQLKEIETRFGSFYAGEDINSSIIKRSNNRSSNKT LINVK.ILK

10

RG2T polynucleotide sequence (SEQ ID NO:126)

GGAAGACGACAATGGTGCAACGGTTGAAGAAGGTTGTGAAAGATAAGAAG
ATGTTCCATTATATTGTCGAGGTGGTTGTAGGGGCAAACACTGACCCCAT
TGCTATCCAGGATACTGTTGCAGATTACCTCAGCATAGAACTGAAAGGAA
15 ATACGAGAGATGCAAGGGCTTATAAGCTTCGTGAATGCTTTAAGGCCCTC
TCTGGTGGAGGTAAGATGAAGTTCCTAGTAATTCTTGACGATGTATGGAG
CCCTGTTGATCTGGATGATATCGGTTTAAGTTCTTTGCCAAATCAAGGTG
TTGACTTCAAGGTCTTGCTGACATCACGCAACAGTGATATCTGCATGATG
ATGGGAGCTAGTTTAATTTTCAACCTCAATATGTTAACAGACGAGGAAGC
20 ACATAATTTTTTCCGTCGATACGCAGAAATTTCTTATGATGCTGATCCCG
AGCTTATTAAGATAGGAGAAGCTATTGTAGAGAAATGTGGTGGTTTACCC
ATTGCCATCAAACTATGGCCGTTACTCTTAGAAATAAACGCAAAGATGC
ATGGAAAGATGCACTTTCTCGTTTAGAGCACCGTGACACTCATAATGTTG
TGGCTGATGTTCTTAAATTGAGCTACAGCAATATCCAAGACGAGGAGACT
25 CGGTCGATTTTTTTGCTATGTGGTTTGTTCCTGAAGACTTTGATATTCC
TACCGAAGACTTAGTGAGGTATGGATGGGGATTGAAAATATTTACCAGAG
TGTATACTATGAGACATGCAAGAAAAAGGTTGGACACGTGCATTGAGCGG
CTTATGCATGCCAACATGTTGATAAAAAGTGATAATGTTGGATTTGTCAA
GATGCATGATCTGGTTCGTGCTTTTGT TTTGGGCATGTTATCTGAAGTCG
30 AGCATGCATCAATTGTCAACCATGGGGATATGCCAGGGTGGTTTGAAACT
GCAAATGATAAGAACAGCTTGTGCAAAAAGAAATTCATTAACATGCAAAGG
TATGTCTGCGATTCTCTGAAGACCTCACGTTTCCAAACCTCTCGATCCTGA
AATTAATGGATGGAGACGAGTCACTGAGGTTTCCTGAAGGCTTTTATGGA
GAAATGGAAAACCTTCAGGTTATATCATATGATAACATGAAGCAGCCATT
35 TCTTCCACAATCACTTCAATGCTCCAATGTTTCGAGTGCTTCATCTCCATC
ACTGCTCATTAAATGTTTGATTGCTCTTCTATTGGAAATCTTTGAATCTC
GAGGTGCTCAGCATTGCTAATTCTGCCATTAAATTGTTACCCTCCACTAT
TGGAGATCTGAAGAAGCTAAGGCTCCTGGATTTGACAAATTGTGTTGGTC
TCTGTATAGCTAATGGCGTCTTTAGAAATTTGGTCAAACCTGAAGAGCTT
40 TATATGAGAGTTGATGATCGAGATTCGTTTTTTGTGAAAGCTGATGACAG
CAAGACCATTACCT

RG2T deduced polypeptide sequence (SEQ ID NO:127)

KTMTMVQRLKKVVKDKKMFHYIVEVVVGANTDPIAIQDTVADYLSIELKGNTRDAR
AYKLRECFKALSGGGKMKFLVILDDVWSPVDLDDIGLSSLPNQGVDFKVLLTSRNS
DICMMMGASLIFNLNMLTDEEAHNFFRRYAEISYDADPELIKIGEAIVEKCGGLPIAI
5 KTMAVTLRNKRKDAWKDALSRLEHRDTHNVVADVLKLSYSNIQDEETRSIFLLCG
LFPEDFDIPTEDLVRYGWGLKIFTRVYTMRHARKRLDTCIERLMHANMLIKSDNVG
FVKMHDLVRAFVLGMLSEVEHASIVNHGDMPGWFETANDKNSLCKRISLTCKGMS
AIPEDLTFPNLSILKLMDGDESLRPEGFYGEMENLQVISYDNMKQPFLPQSLQCSN
VRVLHLHHCSLMFDCSSIGNLLNLEVLSIANSIAIKLLPSTIGDLKKLRLLDLTNCVGL
10 CIANGVFRNLVKLEELYMRVDDRDSFFVKADDSKTTT

RG2U polynucleotide sequence (SEQ ID NO:128)

GCCTTGTGTGGGATGGGTGGAGTGGGAAAGACCACTGTGATGAAGAAGCT
GAAGGAGGTTGTGGTAGGAAAGAAACTGTTTAATCATTATGTTGAGGCGG
15 TTATAGGGGAAAAGACAGACCCCATTTGCTATTCAACAAGCTGTTGCCGAG
TACCTTGGTATAAGTCTAACCGAAACCACTAAACCAGCAAGAACTGATAA
GCTCCGTACATGGTTTGCAAACAACCTCAAATGGAGGAAAGAAGAAGTTCC
TGGTAATACTAGACGATGTATGGCAACCAGTTGATTTGGAAGATATTGGT
TTAAGTCGTTTTCCAAATCAAGATGTTGACTTCAAGGTCTTGATTACATC
20 ACGGGACCAATCAGTTTGCCTGAGATGGGAGTTAAAGCTGATTTAGTTC
TCAAGGTGAGTGTCTTGGAGGAAGCGGAAGCACACAGTTTGTTCCTCCAA
TTTTTAGAACCTTCTGATGATGTCGATCCTGAGCTCAATAAAAATCGGAGA
AGAAATTGTAAAGAAGTGTTCGAGACTACCCATTGCTATCAAAACCATGG
CCTGAACCTCTTAGAAGTAAAAGTAAGGATACATGGAAGAATGCCCTTTCT
25 CGTTTACAACACCATGACATTAACACAATTGCGTCTACTGTTTTCCAAAC
TAGCTATGACAATCTCGAAGACGAGGTGACTAAAGCTACTTTTTTGCTTT
GTGGTTTATTTCCGGAGGACTTCAATATTCTTACCGAGGACCTATTGAGG
TATGGATGGGGATTGAAGTTATTCAAGGAAGTAGATACTATACGAGAAGC
AAGATCCAAGTTGAAAGCCTGCATTGAGCGGCTCATGCATACCAATTTGT
30 TGATCGAAGGTGATGATGTTAGGTACGTAAAGATGCATGATCTGGTGCGT
GCTTTTGTGTTTGGATATGTTTTCTAAAGCCGAGCATGCATCTATTGTCAA
CCATGGTAGTAGTAAGCCAAGGTGGCCTGAAACTGAAAGTGATGTGAGCT
CCTCTTGCAAAGAATTCATTAACATGCAAGGGTNTG

RG2U deduced polypeptide sequence (SEQ ID NO:129)

ALCGMGGVGKTTVMKKLKEVVVGKKLFNHYVEAVIGEKTDPPIAIQQA VAEYLGIS
LTETTKPARTDKLRTWFA NNSNGGKKKFLVILDDVWQPVDLEDIGLSRFPNQDVD
FKVLITSRDQSVCTEMGVKADLV LKVSVEEA EAHSLFLQFLEPSDDVDPELNKIGE
EIVKKCCRLPIAIKTMA.TLR SKSKDTWK NALSRLQHHDINTIASTVFQTSYDNLEDE
40 VTKATFLLCGLFPEDFNIPTEDLLRYGWGLKLFKEVD TIREARSKLKACIERLMHTN

LLIEGDDVRYVKMHDLVRAFLVDMFSKAEHASIVNHGSSKPRWPETESDVSSSKR-
ISLTCKG?

RG2V polynucleotide sequence (SEQ ID NO:130)

5 CTGTGGAAGACACGAATGATSAAGAAGCTGAAGGAGGTCGTGGAACAAAA
GAAAATGTTCAATATTATTGTTCAAGTGGTCATAGGAGAGAAGACAAACC
CTATTGCTATTCAGCAAGCTGTAGCAGATTACCTCTCTATTGAGCTGAAA
GAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTNAATGGTTCGAGGA
CGATGGAGGAAAGAATAAGTTCCTTGTAATACTTGATGATGTATGGCAGT
10 TTGTCGATCTTGAAGATATTGGTTTAAGTCCTCTGCCAAATAAAGGTGTC
AACTTCAAGGTCTTGTGACGTTAAGAGATTCACATGTTTGCACTCTGAT
GGGAGCTGAAGCCAATTCAATTCTCAATATAAAAAGTTTTAAAAGATGTTN
AAGGACAAAGTTTGTTCGCCAGTTTGCTAAAAATGCAGGTGATGATGAC
CTGGATCCTGCTTTCAATGGGATAGCAGATAGTATTGCAAGTAGATGTCA
15 AGGTTTGCCCATGCCATCAAAACCATTGCCTTAAGTCTTAAAGGTAGAA
GCAAGCCTGCGTGGGACCATGCGCTTTCTCGTTTGGAGAACCATAAGATT
GGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTATGACAATCT
CCAAGATGAGGTTACTAAATCTATTTTTWTACTTTGTGCTTTATTTCTG
AAGATTTTGATATTCCTATTGAGGAGTTGGTGAGGTATGGGTGGGGCTTG
20 AAATTATTTATAGAAGCAAAAACCTATAAGAGAAGCAAGAAACAGGCTCAA
CACCTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGGAAGTGATG
ACATTGGATGCGTCAAGATGCACGATGTGGTGCGTGATTTTGTTTGGTAT
ATATTCTCAGAAGTCCAGCACGCTTCAATTGTCAACCATGGTAATGTGTC
AGAGTGGCTAGAGGAAAATCATAGCATCTACTCTTGTAAGAATTTTCAT
25 TAACATGCAAGGGTATGTCTGAGTTTCCCAAAGACCTCAAATTTCCAAAC
CTTTCAATTTTGAACTTATGCATGGAGATAAGTCGNTGAGCTTTCTCTGA
AGACTTTTATGGAAAGATGGAAAAGGTTTCAAGTAATATCATATGATAAAT
TGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACTAACGTTCTGA
GTGCTTCATCTCCATTATTGTTTCAATTAAGGATGTTTGATTGCTCTTCAAT
30 TGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAACATTG
AATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGAT
TTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTAAAAAATTT
GGTCAAACCTGAAGAGCTTTATATGGGTGTTAATGTCCGTATGGACCAGG
CCGT

35

RG2V deduced polypeptide sequence (SEQ ID NO:131)

LWKTRM?KKLKEVVEQKKMFNIIVQVVIGEKTNPPIAQQAVADYLSIELKENTKEAR
ADKLR?WFEDDGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTLRDSH
VCTLMGAEANSILNIKVLKDV?GQSLFRQFAKNAGDDDLDPAFNGIADSIASRCQGL
40 PIAIKTIALSLKGRSKPAWDHALSRLNHNKIGSEEVVREVFKISYDNLQDEVTKSIF?L
CALFPEDFDIPEELVRYGWLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIG

CVKMHADVVRDFVWYIFSEVQHASIVNHG NVSEWLEENHSIYSCKRISLTCKGMSEF
PKDLKFPNLSILKLMHGDKS?SFPEDFYGKMEKVQVISYDKLMYPLLPSSLECSTNV
RVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKG
LRIDNGVLKNLVKLEELYMGVNVVRMDQAV

5

RG2W polynucleotide sequence (SEQ ID NO:132)

TTGGGAAAGAGACAATGATGAAGAATTGAAAGAGGTTGTGGTTGAAAAGA
AAATGTTTAATCATTATGTGGAGGCGTTATAGGGGAGAAGACGGACCCC
ATTGCTATT CAGCAAGCCGTTGCAGAGTACCTTGGTATAATTCTAACAGA
10 AACC ACTAAGGCAGCAAGAACCGATAAGCTACGTGCATGGCTTTCTGACA
ATTCAGATGGAGGAAGAAAGAAGTTCCTAGTAATACTAGACGATGTATGG
CATCCGTTGATATGGAAGATATTGGTTTAAGTCGTTTCCCAAATCAAGG
TGTCGACTTCAAGGTCTTGATTACATCACGGGACCAAGCTGTTTGCCTG
AGATGGGAGTTAAAGCTGATT CAGTTATCAAGGTGAGTGTCTAGAGGAA
15 GCTGAAGCACAAAGCTTATTCTGCCAACTTTGGGAACCTTCTGATGATGT
CGATCCTGAGCTCCATCAGATTGGAGAAGAAATTGTAAGGAAGTGTTGTG
GTTTACCCATTGCAATAAAAACCATGGCCTGCACTCTTAGAAGTAAAAGC
AAGGATACATGGAAGAATGCACTTTCTCGTTTACAACACCATGACATTAA
CACAGTCGCGCCTACTGTTTTTCAAACCAGCTATGACAATCTCCAAGATG
20 AGGTGACTGGAGATACTTTTTTGCTATGTGGTTTGTTCGGAGGACTTC
GATATTCTACTGAAGACTTATTGAAGTATGGATGGGGCTTAAAATTATT
CAAGGGAGTGGATTCTGTAAGAGAAGCAAGATACCAGTTGAACGCCTGCA
TTGAGCGGCTCGTGCATACCAATTTGTTGATTGAAAGTGATGTTGTTGGG
TGCGTCAAGTTGCACGATCTGGTGCCTGCTTTATTTTGGATATGTTTTG
25 TAAAGCGGAGCATGCTTCGATTGTCAACCATGGTAGTAGTAAGCCTGGGT
GGCCTGAAACTGAAAATGATGTGATCAGGACCTCCTGCAAAAGAATCTCA
TTAACATGCAAGGGTATGATTGAGTTTTCTAGTGACCTCAAGTTTCCAAA
TGTCTTGATTTTAAACTTATGCATGGAGATAAGTCGCTAAGGTTT

30 **RG2W deduced polypeptide sequence (SEQ ID NO:133)**

WERDNDEELKEVVVEKKMFNHYVEAVIGEKTDP IAIQQAVA EYLGILTETTKAAR
TDKLRAWLSDNSDGGRRKKFLVILDDVWHPVDMEDIGLSRFPNQGVDFKVLITSRD
QAVCTEMGVKADSVIKVSVLEEAEASLFCQLWEP SDDVDPELHQIGEEIVRKCCG
LPLAIKTMACTLRSSKSDTWKNALSRLQHHDINTVAPT VFQTSYDNLQDEVTGDTF
35 LLCGLFPEDFDIPTEDLLKYGWGLKLFKGVDSVREARYQLNACIERLVHTNLLIESD
VVGCVKLHDLVRAFILDMFCKAEHASIVNHGSSKPGWPETENDVIRT SCKRISLTCK
GMIEFSSDLKFPNVLILKLMHGDKSLRF

RG5 polynucleotide sequence (SEQ ID NO:134)

40 GGGGGGGTGGGGAAGNCGACTCTAGCCCAGAAGNTCTATAATGACCATAA
AATAAAAGGAAGCTTTAGTAAACAAGCATGGATCTGTGTTTCTCAACAAT

ATTCTGATATTTTCAGTTTTGAAAGAAGTCCTTCGGAACATCGGTGTTGAT
TATAAGCATGATGAAACTGTTGGAGAACTTAGCAGAAGGCTTGCAATAGC
TGTCGAAAATGCAAGTTTCTTTCTTGTGTTGGATGATATTTGGCAACATG
AGGTGTGGACTAATTTACTCAGAGCCCCATTAAACACTGCAGCTACAGGA
5 ATAATTCTAGTAACAACTCGTAATGATACAGTTGCACGAGCAATTGGGGT
GGAAGATATTCATCGAGTAGAATTGATGTCAGATGAAGTAGGATGGAAAT
TGCTTTTGAAGAGTATGAACATTAGCAAAGAAAGTGAAGTAGAAAACCTA
CGAGTTTTAGGGGTTGACATTGTCGTTTGTGTGGTGGCCTCCCCCTAGC
CTT

10

RG5 deduced polypeptide sequence (SEQ ID NO:135)

GGVGKTTLAQK?YNDHKIKGSFSKQAWICVSQQYSDISVLKEVLRNIGVDYKHDET
VGELSRRLAIAVENASFLLVLDIWDQHEVWTNLLRAPLNTAATGILVTTRNDTVA
RAIGVEDIHRVELMSDEVGWKLLKSMNISKESEVENLRVLGVLDIVRLCGGLPLAL

15

RG7 polynucleotide sequence (SEQ ID NO:136)

GGTGGGGTTGGGAAGACAACGGGCACAAGGAGGCGACTGCCAATACTTCC
GACTTTTATTCATAGAGATGACGAGTCTTATTTTCTACTACTATAGGGA
GGATATTTGGTTGCGCGAGACGATTCATTGCGCGAAGGGATTCTATCCTT
20 CTTTTTTTCCGCGAAGACTTCGTTCCGGAGGACGGGCTATATTCCCTTTA
ATATTAGTCTAGCCCAGTCTAGGCCAACCATATGGCGATGCGGTAGACCT
CCCAGAGATAGATACTTGATCTTAGAGGATTCACACGTTCAATGGTGGAA
ACTTAAGGAACCGGCTAAGAGTGACTAAACGGAAAAACCCTATTCATTCC
ATAGCCTCATCCGGTCGAGGCATTAAACAATCCATCCCAATCCTCTTTCC
25 TTTGGTCTACTCTAATGATGTGCCCGTTTCGTTGGTGGGAATATCTCTTTAT
ACCGACGATTTATATGGGGATTGCCACTAGCGTTG

30

The above examples are provided to illustrate the invention but not to limit
its scope. Other variants of the invention will be readily apparent to one of ordinary skill
in the art and are encompassed by the appended claims. All publications, patents, and
patent applications cited herein are hereby incorporated by reference.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.
2. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising an leucine rich region (LRR).
3. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising a nucleotide binding site (NBS).
4. The nucleic acid construct of claim 1, wherein the polynucleotide is a full length gene.
5. The nucleic acid construct of claim 1, wherein the further encodes a fusion protein.
6. The nucleic acid construct of claim 1, wherein the RG1 polypeptide is encoded by an RG1 polynucleotide sequence.
7. The nucleic acid construct of claim 6, wherein the RG1 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).
8. The nucleic acid construct of claim 1, wherein the RG2 polypeptide is encoded by an RG2 polynucleotide sequence.
9. The nucleic acid construct of claim 8, wherein the RG2 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of: SEQ ID NO:21 (RG2A);

SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35 (RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

10. The nucleic acid construct of claim 1, wherein the RG3 polypeptide is encoded by an RG3 polynucleotide sequence.

11. The nucleic acid construct of claim 10, wherein the RG3 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:68.

12. The nucleic acid construct of claim 1, wherein the RG4 polypeptide is encoded by an RG4 polynucleotide sequence.

13. The nucleic acid construct of claim 12, wherein the RG4 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:69.

14. The nucleic acid construct of claim 1, wherein the RG5 polypeptide is encoded by an RG5 polynucleotide sequence.

15. The nucleic acid construct of claim 14, wherein the RG5 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:134.

16. The nucleic acid construct of claim 1, wherein the RG7 polypeptide is encoded by an RG7 polynucleotide sequence.
17. The nucleic acid construct of claim 16, wherein the RG7 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:136.
18. The nucleic acid construct of claim 1, further comprising a promoter operably linked to the RG polynucleotide.
19. The nucleic acid construct of claim 18, wherein the promoter is a plant promoter.
20. The nucleic acid construct of claim 19, wherein the plant promoter is a disease resistance promoter.
21. The nucleic acid construct of claim 19, wherein the plant promoter is a lettuce promoter.
22. The nucleic acid construct of claim 18, wherein the promoter is a constitutive promoter.
23. The nucleic acid construct of claim 18, wherein the promoter is an inducible promoter.
24. The nucleic acid construct of claim 18, wherein the promoter is a tissue-specific promoter.
25. A nucleic acid construct comprising a promoter sequence from an RG gene linked to a heterologous polynucleotide.
26. A transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide.

27. The transgenic plant of claim 26, wherein the plant promoter is a plant promoter.
28. The transgenic plant of claim 26, wherein the plant promoter is a viral promoter.
- 5 29. The transgenic plant of claim 26, wherein the plant promoter is a heterologous promoter.
30. The transgenic plant of claim 26, wherein the plant is lettuce.
- 10 31. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).
- 15 32. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:21 (RG2A); SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35 (RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).
- 20 33. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:68 (RG3) and SEQ ID NO:69 (RG4).
- 25 34. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:70 (RG5) and SEQ ID NO:71 (RG6).
- 30 35. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:72 (RG7) and SEQ ID NO:73 (RG8).

34. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a sequence as set forth in SEQ ID NO:134 (RG5).

35. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a
5 sequence as set forth in SEQ ID NO:136 (RG7).

36. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO: 13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ
10 ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), and SEQ ID NO:20 (RG1J).

37. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41
15 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ
20 ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID
25 NO:133 (RG2W).

38. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG3 polypeptide with a sequence as set forth by SEQ ID NO:138.

39. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG4 polypeptide with a sequence as set forth by SEQ ID NO:139.

40. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135.

41. A method of enhancing disease resistance in a plant, the method comprising
5 introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence.

42. The method of claim 41, wherein the plant is a lettuce plant.

10 43. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51
15 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O);
20 SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

44. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide
25 selected from the group consisting of SEQ ID NO:138 (RG3); SEQ ID NO:139 (RG4); and SEQ ID NO:135 (RG5).

45. The method of claim 41, wherein the promoter is a tissue-specific promoter or a plant disease resistance promoter.

46. The method of claim 41, wherein the promoter is a constitutive promoter or an inducible promoter.

5 47. A method of detecting RG resistance genes in a nucleic acid sample, the method comprising:

contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and,

wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample.

10

48. The method of claim 47, wherein the RG polynucleotide is an RG1 polynucleotide.

49. The method of claim 47, wherein the RG polynucleotide is an RG2 polynucleotide.

15 50. The method of claim 47, wherein the RG polynucleotide is an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide.

51. The method of claim 47, wherein the RG resistance gene is amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

20

52. The method of claim 51, where the RG resistance gene is amplified by the polymerase chain reaction.

53. The method of claim 47, wherein the RG polynucleotide is labeled.

25

54. An RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : 435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARAN et al. Development of Reliable PCR-Based Markers Linked to Downy Mildew Resistance Genes in Lettuce. Theor. Appl. Genet. 1993. Vol. 85, No. 8, pages 985-993, see entire article.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	KESSELI et al. Analysis of a Detailed Genetic Linkage Map of Lactuca sativa (Lettuce) Constructed From RFLP and RAPD Markers. Genetics. April 1994. Vol. 136, No. 4, pages 1435-1446, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	MICHELMORE, RW. Isolation of Disease Resistance Genes from Crop Plants. Current Opinion in Biotechnology. 1995. Vol. 6, No. 2, pages 145-152, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:		*T*	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A	document defining the general state of the art which is not considered to be of particular relevance	*X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B	earlier document published on or after the international filing date	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A*	document member of the same patent family
O	document referring to an oral disclosure, use, exhibition or other means		
P	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

14 MARCH 1998

Date of mailing of the international search report

13 APR 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized official

PHUONG BUI

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARAN et al. Recent Amplification of Triose Phosphate Isomerase Related Sequences in Lettuce. Genome. 1992. Vol. 35, No. 4, pages 627-635, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	PARAN et al. Identification of Restriction Fragment Length Polymorphism and Random Amplified Polymorphic DNA markers linked to Downy Mildew Resistance Genes in Lettuce, Using Near-Isogenic Lines. Genome. 1991. Vol. 34, No. 6, pages 1021-1027, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US98/00615

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 7, 9, 11, 13, 15, 17, 31-40, 43-44
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

these claims are drawn to numerous sequences identified by SEQ ID NOs. However, since no computer readable form was submitted, no meaningful search could be carried out.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

A01H 1/00; C07H 21/04; C07K 14/00; C12N 5/04, 5/10; C12P 19/34; C12Q 1/68

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☒ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)